

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 23:09:34 ; Search time 6466 Seconds

(without alignments)
10723.547 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 1482
Sequence: 1 CAGACACACAAAGCGATCT.....AATATTTCGAACTATTA 1482Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hcc.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_ges1.*
10: gb_ges2.*
11: gb_ges3.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	291.4	19.7	317	10	T09652 0224m3 gmbp
2	256	17.3	1721	8	DR145459 49056454
3	249.4	16.8	1808	8	DR125630 49096520
4	248.4	16.8	1519	10	AG386893 Mus muscu
5	247.2	16.7	1453	10	AJ591978 Arabidops
6	244.4	16.5	2071	8	DR147304 49027161
7	243	16.4	1932	8	DR141720 49300102
8	241	16.3	1727	10	CL114085 ISB1-59J1
9	240	16.2	1491	10	CG753221 P048-2-A0
10	238.4	16.1	1783	8	DR141824 49190649
11	237.8	16.0	1434	10	AJ592058 Arabidops
12	237.8	16.0	1808	10	CL078963 CH216-154
13	235	15.9	1654	10	CL118709 ISB1-7211
14	234.2	15.8	1797	10	CL066150 CH216-108
15	234	15.8	1829	10	CL109409 ISB1-51H2
16	233.8	15.8	1736	10	CL066373 CH216-108
17	233.6	15.8	1407	10	AJ592026 Arabidops
18	233.6	15.8	1753	8	DR131208 49217594
19	233.8	15.7	1512	10	CL113943 ISB1-59E2
20	233.8	15.7	1531	10	CG748014 P041-4-B0
21	232.6	15.7	1811	10	CG753732 P048-4-G0
22	231.6	15.6	1917	10	CL081899 CH216-164

23	231.4	15.6	1650	10	CL041163	CL041163 CH216-51P
24	231.2	15.6	1762	10	CL082587	CL082587 CH216-168
25	230.4	15.5	1824	10	CL081234	CL081234 CH216-161
26	230.2	15.5	1727	10	CL118815	CL118815 ISB1-7205
27	230	15.5	1632	10	CL082569	CL082569 CH216-167
28	229.8	15.5	1616	10	CL081995	CL081995 CH216-165
29	229.6	15.5	1630	10	CL078686	CL078686 CH216-152
30	229.2	15.5	1457	10	CL082658	CL082658 CH216-169
31	228.8	15.4	1566	10	CG757757	CG757757 P053-1-D0
32	228.6	15.4	1434	9	CC187638	CC187638 CH261-98P
33	228.4	15.4	1507	10	AG346189	AG346189 Mus muscu
34	228.4	15.4	1823	10	CL058635	CL058635 CH216-188H
35	228	15.4	1259	11	CR004722	CR004722 Forward s
36	228	15.4	2033	10	CL081977	CL081977 CH216-165
37	228	15.4	2103	10	CL041093	CL041093 CH216-51M
38	227.6	15.4	1712	10	CL078539	CL078539 CH216-151
39	226.8	15.3	1547	10	CL066420	CL066420 CH216-108
40	226.6	15.3	1377	10	CL077307	CL077307 CH216-144
41	226.4	15.3	1716	9	CC222065	CC222065 CH261-11A
42	225.6	15.2	1562	10	CL078533	CL078533 CH216-151
43	225.6	15.2	1651	10	CL079030	CL079030 CH216-154
44	225.4	15.2	1482	10	AG332292	AG332292 Mus muscu
45	224.2	15.1	1506	10	CL082761	CL082761 CH216-171

ALIGNMENTS

RESULT 1
T09652 317 bp DNA linear GSS 05-DEC-2000
LOCUS 0224m3 gmbpFHB3.1, G. Roman Reddy Plasmodium falciparum genomic
DEFINITION clone 0224m, genomic survey sequence.
ACCESSION T09652
VERSION T09652.1 GI:319484
KEYWORDS GSS.

ORGANISM Plasmodium falciparum (malaria parasite P. falciparum)

Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 317)
Reddy,G.R., Chakrabarti,D., Schuster,S.M., Ferl,R.J., Almira,E.C.
and Dame,J.B.TITLE Gene sequence tags from Plasmodium falciparum genomic DNA fragments
prepared by the genease activity of mung bean nuclease

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90, 9867-9871 (1993)

PUBMED 8234327

COMMENT Other GSSs: 0224m7

Contact: Dame JB

Dept. of Pathobiology, College of Veterinary Medicine

University of Florida

2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA

Tel: 352 392 4700

Fax: 352 392 9704

Email: damej@mail.ufl.edu

Seq primer: T3

Class: Shotgun.

FEATURES

source

Location/Qualifiers

1. 317

/organism="Plasmodium falciparum"

/mol_type="genomic DNA"

/db_xref="taxon:5833"

/clone="0224m"

/lab_host="E. coli XL1-Blue"

/clone_id="gmbpFHB3.1, G. Roman Reddy"

/note="Vector: pBluescript SK(+); Genomic DNA, from
asynchronous blood stage parasites of the cloned Honduran
HB3 isolate cultured in vitro, was digested with mung bean
nuclease in the presence of 30% formamide at 50C
(Vernick, K.D., Imberiski, R.B., and McCutchan, T.F. 1988.
Nucleic Acids Research 16:6883-6896). The ends of the
fragments were polished using T4 DNA polymerase, and the
fragments were ligated to EcoR V-cleaved and
dephosphorylated pBluescript SK(+). Recombinant plasmids

Oy		902	GACATCTTGAAGAAAGGAAAAGAATGGTTCAATAAACCAGAACAAAAAGGATTAATCTG	961
Db		1037	AA	1096
Oy		962	CTGCATACACAATAATCATTCATTAGAGACAGTAATAATTTCTGATGTTAATGATTTTCAA	1021
Db		1097	AA	1156
Oy		1022	TAACTAAGTATGAGAGTAAATTAAGTGCTGAATATGACGATTCAATTAATAGATGAGAG	1081
Db		1157	AA	1216
Oy		1082	AAGATGATGAGACTTAGACAAATTTTAAGCCTATTGTCGCAATATGACAAATTTCCAGATG	1141
Db		1217	AA	1276
Oy		1142	AAGAAAAATCATGGAATTTATTAAGAACTAGAGAATTTGATGAGAAAAATGAAATTTAG	1201
Db		1277	AA	1336
Oy		1202	ATGATTTAGATGAGAAATAGAAAAATCATCAGAAATTAATCTGAAGAAAAATTAATA	1261
Db		1337	AA	1396
Oy		1262	AAGGAAAGAAATATGAAAAAAACAAGGATTAATTTTAAACCAATGATTAAGTTGT	1321
Db		1397	AA	1456
Oy		1322	ATGATGAGCAATTAATAATTAATAATGATTAAGCAGTTAATATAGAGAAAGGAAAT	1381
Db		1457	AA	1516
Oy		1382	TCATATAA 1389 	
Db		1517	AAAAAAAA 1524	
RESULT 3				
DR125630				
LOCUS				
DEFINITION			DR125630 1808 bp mRNA linear EST 15-JUN-2005	
ACCESSION			49095620 Drosophila pseudoobscura embryonic cDNA library drosophilala	
VERSION			pseudobsecura cDNA clone O23 5', mRNA sequence.	
KEYWORDS			DR125630	
SOURCE			DR125630..1 GI:67842328	
ORGANISM			EST.	
			Drosophila pseudoobscura	
			Drosophila pseudoobscura	
			Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
			Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
			Ephyridoidae; Drosophilidae; Drosophila.	
REFERENCE			1 (bases 1 to 1808)	
AUTHORS			Richards, S., Liu, Y., Bettencourt, B. R., Hrabecky, P., Letovsky, S., Nielsen, R., Thornton, K., Hubisz, M. J., Chen, R., Meisel, R. P., CoRoone, O., Hua, S., Smith, M. A., Zhang, P., Liu, J., Bussemaker, H. J., van Batenburg, M. F., Howells, S. L., Scherer, S. E., Sodergren, E., Matthews, B. B., Crosby, M. A., Schroeder, A. J., Ortiz-Barrioneto, D., Rives, C. M., Metzker, M. L., Muzny, D. M., Scott, G., Steffen, D., Wheeler, D. A., Worley, K. C., Havlak, P., Durbin, K. J., Ban, A., Gill, R., Hume, J., Morgan, M. B., Miner, G., Hamilton, C., Huang, Y., Waldron, L., Verdusco, D., Cleic-Blankenbury, K. P., Dubchak, I., Noor, M. A., Anderson, W., White, K. P., Clark, A. G., Schaeffer, S. W., Gelbart, W., Weinstein, G. M. and Gibbs, R. A. Comparative genome sequencing of Drosophila pseudoobscura: Chromosome1, gene, and cis-element evolution Genome Res. 15 (1), 1-18 (2005) 15632085	
JOURNAL			Contact: Stephen Richards	
PUBMED			Human Genome Sequencing Center	
COMMENT			Baylor College of Medicine	
			One Baylor Plaza, Houston, TX 77030, USA	
			Tel.: 713-798-6667	
			Email: stephenr@bcm.tmc.edu	
			NCBI Trace Archive: 226697953	

FEATURES		Insert Length: 1750	Std Error: 0.25
SOURCE		Location/Qualifiers	
1. .1808			
		/organism="Drosophila pseudoobscura"	
		/mol_type="mRNA"	
		/db_xref="taxon:7237"	
		/clone="023"	
		/dev_stage="0-18h embryos"	
		/clone_id="Drosophila pseudoobscura embryonic cdna library"	
		/note="Vector: pOT87; Site_1: EcoRI; Site_2: XhoI; oligo dt priming from poly A+ RNA, directionally cloned"	
ORIGIN			
Query Match	16.8%	Score 249.4;	DB 8; Length 1808;
Best Local Similarity	48.6%	Pred. No. 3.1e-31;	
Matches 670;	Conservative 0;	Mismatches 709;	Indels 0; Gaps 0;
QY	2	AAAGAACAAACGCGATCTTGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAC	61
DB	340	AAAGAA	399
QY	62	AAAGCGATTTAGACAGATAGACTTGCTTAAGAAAAGTTACAAAGACAGCAAGCGATT	121
DB	400	AA	459
QY	122	TAGAACAGAGAGACTTGCTTAAGAAAAGTTGCAAGAACAAACGACTTACACAG	181
DB	460	AAAAAAAAAAAAAAAAAGAAAGAA	519
QY	182	AGACGCTGCTTAAGAAAAGTTGCAAGAACAAACGACTTGAACAGAGAGAGCTG	241
DB	520	AAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAAAAAAAGAA	579
QY	242	CTTAAGAAAAGTTGCAAGAACAAACGACTTGAACAGATAGACTTGCTTAAGAAA	301
DB	580	AAAAAAAAAGAAAGAAA	639
QY	302	AGTTACAGAGCAGCAAGCGATTTGAACAGAGAGCGTCTTAAGAAAAGTTGCAAG	361
DB	640	AAAGAAAAAAAAAGAAAAAG	699
QY	362	AACACAAAGCGATTTGACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAACAA	421
DB	700	AAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAA	759
QY	422	GCGATTTGAACAGAGAGACTTGCTTAAGAAAAGTTGCAAGAACAAACGACTTTAG	481
DB	760	AAAAAAAAAGAAA	819
QY	482	AACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAACGACTTTGAACAGAGAG	541
DB	820	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAAGAAAGAAAAAAAA	879
QY	542	GACGTCTTAAGAAAAGTTGCAAGAACAAACGACTTTGAACAGAGAGCGTCTGA	601
DB	880	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAGAA	939
QY	602	AAAGAAAGTTGCAAGACAGCAAGATTTGAACAAAGGAAGCGTATCGAAAAAAA	661
DB	940	AAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAA	999
QY	662	ATTAGAAAGAAAGGAACATGAGATATATTAGCAGAGATTTATATGTCGTTAG	721
DB	1000	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAA	1059
QY	722	AAATACCACTATGAACCTTCATCAGAAAAATGAACGTGATATTATATCCACATCAAT	781
DB	1060	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAA	1119
QY	782	CTTCTTTATCTCAGAGCAACGAGGAAATGTGAGATTCAGAGAAATATCTTAATAG	841
DB	1120	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAAGAAAAAAAAAGAAAAAA	1179

[illegible]

COMMENT	FEATURES	ORIGIN
<p>Gaston Cremieux, 91057 Evry cedex, FRANCE</p> <p>PCR was performed on DNA from transformants of <i>Arabidopsis thaliana</i> plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at http://dbsgap.versailles.inra.fr/publiclines/. This sequence has been generated in the framework of the French plant genomics program 'Géoplanté' (http://www.geoplantec.com and http://geoplantec-info.infobiogen.fr).</p>	<p>Location/Qualifiers</p> <p>1..1453</p> <p>/organism="Arabidopsis thaliana"</p> <p>/mol_type="genomic DNA"</p> <p>/db_xref="taxon:3702"</p> <p>/clone="598D01"</p> <p>/clone_lib="Arabidopsis thaliana T-DNA insertion lines"</p> <p>/ecotype="Wassilewskija"</p> <p>1..1453</p> <p>/note="T-DNA flanking sequence</p> <p>right border"</p>	<p>Query Match 16.7%; Score 247.2; DB 10; Length 1453;</p> <p>Best Local Similarity 44.0%; Pred. No. 7.2e-31;</p> <p>Matches 618; Conservative 0; Mismatches 785; Indels 0; Gaps 0;</p>
<p>2 AAGAACAAACAAGCGATCTAGACAAGAGAGACGCTGTAAGAAAGTTGCCAAGACAC 61</p> <p>1419 AA 1366</p>		
<p>62 AAAGCGATTTGAAACAAGATGACTCTGTAAGAAAAATTACAAGACGACAAACGATT 121</p> <p>1359 AA 1300</p>		
<p>122 TAGACAAGAGAGACTTGCTTAAGAAAAGTTGCAAGAACAAAGCATTTAGACAG 181</p> <p>1299 AA 1240</p>		
<p>182 AGAGCGCTGCTAAAGAAAGTTGCAGAACAAAGCGATTTAGAACAGAGAGACGTG 241</p> <p>1239 AA 1180</p>		
<p>242 CTAAAGAAAGTTGCCAAGAACAAAGACGATTTAGAACAGATGACTGCTAAGAA 301</p> <p>1179 AA 1120</p>		
<p>302 AGTTACAAAGACGACAAACGATTTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAG 361</p> <p>1119 AA 1060</p>		
<p>362 AACACAAAGCGATTTGAAACAAGAGAGACGTGCTAAAGAAAAGTTGCCAAGACAA 421</p> <p>1059 AA 1000</p>		
<p>422 GCGATTTAGAACAGAGAGACTTGCTTAAGAAAAGTTGCAGAACAAAGCGATTTAG 481</p> <p>999 AA 940</p>		
<p>482 AACACAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGA 541</p> <p>939 AA 880</p>		
<p>542 GACGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGACGTGCT 601</p> <p>879 AA 820</p>		
<p>602 AAGAAAAGTTGCCAAGACGACAAAGAGATTTAGAACAAAGAGCGCTGATCGAAAAA 661</p> <p>819 AA 760</p>		

QY	662	ATTTTGAAAGAAAAAGGAAACATGAGATTAATTAAGCAGAGATTATATGTCGTTAG	721
Db	759	AAA	700
OY	722	AAATACGACGTATGAACTTCATCAGAAAAATGAAACGATATTATATACCATCAAT	781
Db	699	AAA	640
OY	782	CTTCTTTACCTCAGACAAACAGAGGAATAGAGATTCCAGAAATATCTAATAG	841
Db	639	AAA	580
OY	842	AAAAAACCAATAGAGATCTTATTCACAAATGTTGAAAGACGAAGGATTAATAG	901
Db	579	AAA	520
OY	902	GACATCTTGAAAGAAAGAGATGGTTCATTAATAACGAACAAAGAAATATCTG	961
Db	519	AAA	460
OY	962	CTGACATCAAAATCATATCATTTAGACAGATAATTTCTGATGTTAATGATTTCAA	1021
Db	459	AAA	400
OY	1022	TAACTAGTATGAGATGAATTAAGTCTGTAATATGACATTCATTAATAGTAGAG	1081
Db	399	AAA	340
OY	1082	AAGATGATGAAGACTTACAGCAATTTAGCCTATTGTGCATATGACATTTCCAGATG	1141
Db	339	NAAA	280
OY	1142	AAGAAACATGAGAAATTTATTAAGAACTAGAAATTTGATGAGAAAAATGAAATTTAG	1201
Db	279	AAA	220
OY	1202	ATGATTTAGTAGAGCAATTCAGAAAAATCATCAGAAATTTCTGAGAAAAATTAATA	1261
Db	219	AAA	160
OY	1262	AAGGAAAGAAATATGAAAAACAAGTAATATTTAAACCAATGATTAAGTTGT	1321
Db	159	AAA	100
OY	1332	ATGATGACATATTTAAAAATATTAATAATGATTAAGCAGGTTAATAGCAAAAGAAAT	1391
Db	99	AAA	40
OY	1392	TCATTAATCATTTGTTCAATATA	1404
Db	39	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	17
RESULT 6			
DRI47304			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			

TITLE	JOURNAL	COMMENT
Wheeler,D.A., Worley,K.C., Havlik,P., Durbin,K.J., Egan,A., Gill,R., Hume,J., Morgan,M.B., Miner,G., Hamilton,C., Huang,Y., Waldron,L., Verdusco,D., Cleric-Bianenburg,K.P., Dubchak,I., Noor,M.A., Anderson,W., White,K.P., Clark,A.G., Schaeffer,S.W., Gelbart,W., Weinstein,G.M. and Gibbs,R.A.	Comparative genome sequencing of Drosophila pseudoobscura: chromosomal, gene, and cis-element evolution	Genome Res. 15 (1), 1-18 (2005)
15632085	Contact: Stephen Richards	Human Genome Sequencing Center
	Baylor College of Medicine	One Baylor Plaza, Houston, TX 77030, USA
	Tel: 713-798-6667	Email: stephen@bcm.tmc.edu
	NCBI Trace Archive: 226691688	Insert Length: 1750 Std Error: 0.25.
FEATURES	location/Qualifiers	
source	1..2071	
	/organism="Drosophila pseudoobscura"	
	/mol_type="mRNA"	
	/db_xref="taxon:7237"	
	/clone="O11"	
	/dev_stage="0-18h embryos"	
	/clone_lib="Drosophila pseudoobscura embryonic cDNA library"	
	/note="Vector: pOTB7, Site_1: EcoRI; Site_2: XhoI; oligo dt priming from poly A+ RNA, directionally cloned"	
ORIGIN		
Query Match	16.5%; Score 244.4; DB 8; Length 2071;	
Best Local Similarity	47.7%; Pred. No. 2.1e-30;	
Matches 707; Conservative	0; Mismatches 774; Indels 0; Gaps 0	
QY	2 AAGAACAAACAAGCGATCTTAGAACACAGACGCTGCTTAAGAAAGTTGCAAGAACAC	61
DB	492 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAGAAAAAAAAAGAAAAA	551
QY	62 AAAGCGATTAGAACAGATGACTGCTAAGAAAAAGTTACAGAGCACAACCGATT	121
DB	552 GAAAGAAAGAAAGAAAAAAAAAAAAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAAA	611
QY	122 TAGAACAGAGACTGCTTAAGAAAGTTGCAAGAACAAACGATCTTAGAACAG	181
DB	612 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAGAAAAAAAAAGAAAAA	671
QY	182 AGAGCGCTGAAGAAAGTTGCAAGAACAAACGATTTAGAACAGAGACGTG	241
DB	672 AAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAGAAAAAAAAAGAAAA	731
QY	242 CTAAAGAAAGTTGCAAGAACAAACGATTTAGAACAGATGACTGCTAAGAA	301
DB	732 AAAAAAAAAAGAAAGAAAAAAAAAAAAAAAAAGAAAGAAAGAAAAAAAAAGAAAA	791
QY	302 AGTTCAAGACGACGAAGCGATTTAGAACAGAGACGCTGCTTAAGAAAGTTGCAAG	361
DB	792 AAAAAAAAAAAAAAAAAAGAAAGAAAAAAAAAGAAAGAAAGAAAAAAAAAGAAAAA	851
QY	362 AACACAAAGCGATTTAGAACAGAGACGCTGCTAAGAAAGTTGCAAGAACAA	421
DB	852 AAAAAAAAAAGAAAGAAAAAAAAAAAAAAAAAGAAAGAAAGAAAAAAAAAGAAAA	911
QY	422 GCGATTTAGAACAGAGACTGCTAAGAAAGTTGCAAGAACAAACGATTTAG	481
DB	912 AAAAAAAAAAAAAAAAAAGAAAGAAAAAAAAAGAAAGAAAGAAAAAAAAAGAAAA	971
QY	482 AACACAGAGACGCTGCTAAGAAAGTTGCAAGAACAAACGATTTAGAACAGAG	541
DB	972 AAAAAAAAAAGAAAGAAAAAAAAAGAAAGAAAGAAAAAAAAAGAAAAA	1031
QY	542 GACGCTGAAGAAAGTTGCAAGAACAAACGATTTAGAACAGAGACGCTG	601
DB	1032 AAAAAAAAAAAAAAAAAAGAAAGAAAAAAAAAGAAAGAAAGAAAAAAAAAGAAAA	1091

```
QY 602 AAGAAAAGTTGCAAGACGACGAAAGAGATTGAAACAAGAAAGGCTGATACGAAAAA 661
DB 1092 AAAAAAAAAAGAAAAAAAAAGAAAAAAAAAGAAAAAAAAAGAAAAAAAAAGAAAAA 1151
QY 662 ATTTAGAAAGAAAAAGAAACATGAGATATATTAGCAGAGATTTTATGTGCTTAG 721
DB 1152 AAGAAAAAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1211
QY 722 AAATACCAAGCTATAGAACTTCCATCAGAAAAAGAACTGATATTATATACCAATCAAT 781
DB 1212 GAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1271
QY 782 CTCTCTTACCTCAGACAAACAGAGGAAATAGTAGATTCACAGAAATTTCTATATAG 841
DB 1272 AAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1331
QY 842 AAAAAACAATAGAGATTTATTCAACAATGTTGAAGGACGAAAGGATATACATTAAG 901
DB 1332 AAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1391
QY 902 GACATCTTGAAGAAAGAAAGATGCTTCAATTAAGCCAGAAACAAAGAAATTAATCTG 961
DB 1392 GAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1451
QY 962 CTGACATCAAAATCATCATCTAGACAGCTAAATTTCTGATGTTAATGATTTTCAA 1021
DB 1452 AAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1511
QY 1022 TAAGTAAAGTAAAGATGAATTAAGTGTGAATTAAGCAATTAATTAAGTAAAGAG 1081
DB 1512 AAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1571
QY 1082 AAGATGATGAAGACTTGAAGCAATTTAAGCTATTGTGCAATTAAGCAATTTCCAGATG 1141
DB 1572 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1631
QY 1142 AAGAAAACATGCAATTTATTAAGAACTAGAAAGTTTGATAGAGAAAAAGAAATTTAG 1201
DB 1632 AGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGCAAGAAAG 1691
QY 1202 ATGATTTAAGTAAAGAAATGAAGAAATCATCAGAAAGAAATTAATCGAAGAAAAATTA 1261
DB 1692 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1751
QY 1262 AAGCAAAAGAAATGAAGAAACAAAGATTAATTTTAAACCAATGATTAAGATTTGT 1321
DB 1752 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1811
QY 1322 ATGATGACATTAATTAAGAAATATTAAGTAAGCAAGTTAATTAAGCAAGAAAGAAAT 1381
DB 1812 AAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1871
QY 1382 TCATTAATCACTTCTTCATATATTGAAGCAATGAATTTTCAAGTCGTGATG 1441
DB 1872 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1931
QY 1442 AGTTATCTGAAGATATTAATTAATTTTATGAAGATTA 1482
DB 1932 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1972

RESULT 7
DR141720 1932 bp mRNA linear EST 16-JUN-2005
LOCUS DR141720
DEFINITION 49300102 Drosophila pseudoobscura embryonic cDNA library Drosophila
ACCESSION DR141720.1 GI:67886810
VERSION DR141720
KEYWORDS EST.
SOURCE Drosophila pseudoobscura
ORGANISM Drosophila pseudoobscura
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
```

```
REFERENCE
AUTHORS Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1932)
Richards, S., Liu, Y., Bettencourt, B. R., Hradecky, P., Letovsky, S.,
Nielsen, R., Thornton, K., Hubisz, M. J., Chen, R., Weisel, R. P.,
Couronne, O., Hua, S., Smith, M. A., Zhang, P., Liu, J., Bussemaker, H. J.,
van Batenburg, M. F., Howells, S. L., Scherer, S. E., Sodergren, E.,
Matthews, B. B., Crosby, M. A., Schroeder, A. J., Ortiz-Barrientos, D.,
Rives, C. M., Metzker, M. L., Muzny, D. M., Scott, G., Steffen, D.,
Wheeler, D. A., Worley, K. C., Havlak, P., Dudin, K. J., Egan, A.,
Gill, R., Hume, J., Morgan, M. B., Miner, G., Hamilton, C., Huang, Y.,
Maldron, L., Verdusco, D., Clerc-Blankenburg, K. P., Dubchak, I.,
Noor, M. A., Anderson, W., White, K. P., Clark, A. G., Schaeffer, S. W.,
Gelbart, W., Weinstock, G. M., and Gibbs, R. A.
Comparative genome sequencing of Drosophila pseudoobscura:
Chromosomal, gene, and cis-element evolution
Genome Res. 15 (1), 1-18 (2005)
15632085
JOURNAL Contact: Stephen Richards
PUBMED Human Genome Sequencing Center
COMMENT Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephen@bcm.tmc.edu
NCBI Trace Archive: 226717045
Insert Length: 1750 Std Error: 0.25.
FEATURES
source
1..1932
Location/Qualifiers
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/clone="113"
/dev_stage="0-18h embryos"
/clone_lib="Drosophila pseudoobscura embryonic cDNA
library"
/note="Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; oligo
de priming from poly A+ RNA, directionally cloned"
ORIGIN
Query Match 16.4%; Score 243; DB 8; Length 1932;
Best Local Similarity 48.0%; Pred. No. 3.5e-30;
Matches 666; Conservative 0; Mismatches 722; Indels 0; Gaps 0;
QY 2 AAGAACAAAGCGATCTGAACAAAGAGAGCTGTTAAGAAAGTTGCAAGAACAC 61
DB 251 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 310
QY 62 AAGCGATTTAAGAACAAAGATGAGCTTGAAGAAAGTTTACAAAGACCAAGCGATT 121
DB 311 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 370
QY 122 TAGAACAAAGAGACTTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTGAACAG 181
DB 371 AGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 430
QY 182 AGAGACGTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTAAGAACAAAGAGAGCTG 241
DB 431 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 490
QY 242 CTAAGAAAGTTGCAAGAACAAAGCGATTTAAGAACAAAGTAGACTTGCTTAAGAAA 301
DB 491 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 550
QY 302 AGTTACAAAGACGCAAGCGATTTAAGAACAAAGAGAGCTGCTAAAGAAAGTTGCAAG 361
DB 551 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 610
QY 362 AACACAAAGCGATTTAAGAACAAAGAGAGCTGCTAAGAAAGTTGCAAGAACAA 421
DB 611 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 670
QY 422 GCGATTTAAGAACAAAGAGAGCTTGCTAAGAAAGTTGCAAGAACAAAGCGATTTAG 481
```



```

QY 602 AAGAAAGTTGCAAGCAGCAAGAGATTGAAACAAAGAGCGTGTACGAAAAA 661
Db 891 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 832
QY 662 ATTTAGAAAGAAAAAGAACATGAGATATATTAGCAGAGATTATATGCGTTAG 721
Db 831 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 772
QY 722 AATATCCAGCTATGAACTTCATCGAAAAATGAAAGTGGATTTTATCCACATCAT 781
Db 771 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 712
QY 782 CTTCCTTACCTGAGCAAGAGGAGATAGATAGAGATTCCAAGAAATATCTATATAG 841
Db 711 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 652
QY 842 AAAAAACAAATAGAGATCTTATTACAACAATGTTGAGAGAGAGGATATACATAAG 901
Db 651 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 592
QY 902 GACATCTTGAAGAAAGAGATGTTCAATTAACCAAGACAAAAAGATTAATCTG 961
Db 591 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 532
QY 962 CTGACATCAAAATCATATAGAGAGAGTAAATTTCTGATGTTAATGATTTTCAA 1021
Db 531 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 472
QY 1022 TAAGTAAATGATGAGATGAATTAAGTCTGAATATGACGATTCATTAAATGAGAG 1081
Db 471 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 412
QY 1082 AAGATGATGAAGACTTGAAGATTTAAGCTATTTGCAATATGCAATTTCCAGATG 1141
Db 411 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 352
QY 1142 AAGAAACATAGGAATTTTAAAGAACTGAAGATTGATGAGAAAAATGAATTTAG 1201
Db 351 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 292
QY 1202 ATGATTTAGATGAAGATAGAAAAATCATCAGAGAAATTTATCTGAGAAAAATTA 1261
Db 291 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 232
QY 1262 AAGGAAGAAATRTGAAAAACAAAGATTAATTTTAACCAATGTTAAAGTTTGT 1321
Db 231 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 172
QY 1322 ATGATGAGATATTAAATAATTAATAATGATTAAGCAGGTTAATAAGAAAAAT 1381
Db 171 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 112
QY 1382 TCATTAATATCA 1392
Db 111 AAAAAAAAAANNA 101

```

```

RESULT 10
DRI41924 1783 bp mRNA linear EST 16-JUN-2005
LOCUS 4910649 Drosophila pseudoobscura embryonic CDNA library Drosophila
DEFINITION pseudoobscura CDNA clone O7 3', mRNA sequence.
ACCESSION DRI41924
VERSION DRI41924.1 GI:67887014
KEYWORDS EST.
SOURCE Drosophila pseudoobscura
ORGANISM Drosophila pseudoobscura
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1783)
AUTHORS Richardt,S., Liu,Y., Bettencourt,B.R., Hradecky,P., Letovsky,S.,

```

```

TITLE
JOURNAL
PUBMED
COMMENT
Contact: Stephen Richards
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephenr@bcm.tmc.edu
NCBI Trace Archive: 226707030
Insert Length: 1750 Std Error: 0.25.
Location/Qualifiers
1..1783
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/clone="O7"
/dev_stage="0-18h embryos"
/clone_id="Drosophila pseudoobscura embryonic CDNA
library"
/note="Vector: pOTB7, Site 1: EcoRI; Site 2: XhoI; oligo
dt priming from poly A+ RNA, directionally cloned"
ORIGIN
Query Match 16.1%; Score 238.4; DB 8; Length 1783;
Best Local Similarity 48.0%; Pred. No. 2e-29;
Matches 680; Conservative 0; Mismatches 738; Indels 0; Gaps 0;
QY 6 ACAACAAAGCATCTAGAACAGAGAGCGTCTTAAGAAAGTTGCAAGAACACAAAG 65
Db 344 AAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 403
QY 66 CGATTGACAAAGTAGACTTGTCTAAAGAAAAGTTACAGAGCGCAAGCGATTTAGA 125
Db 404 AAAAAAAAAAGAGGAGAAAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 463
QY 126 ACAAGAGACTTGTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGAG 185
Db 464 AAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 523
QY 186 ACGTCTTAAGAAAAAGTTGCAAGAACAAAGCATTTAGACAAAGAGAGCGTCTTA 245
Db 524 AAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 583
QY 246 AAGAAAGTTGCAAGAACAAAGCGATTTGAACAAAGATGACTTGTAAAGAAAAGTT 305
Db 584 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 643
QY 306 ACAAGAGCAAGAACGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGACA 365
Db 644 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 703
QY 366 ACAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGA 425
Db 704 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 763
QY 426 TTTGAACAAGAGACTTGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGACA 485
Db 764 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 823
QY 486 AAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCG 545

```



```

Db      824 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAA 883
Qy      546 TGTAAAGAAAGTTGCAAGACAAACGATTTAGAAACAAGAGACGTCTAAGA 605
Db      884 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAA 943
Qy      606 AAAGTTGCAAGACAGAAAGATTTTGAACAAAGAAAGCGTGTATCGAAAAAATTT 665
Db      944 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAA 1003
Qy      666 AGAAGAGAAAAAGAAACATGAGATATATTAGCAGAGATTTATATGTCGTTAGAAAT 725
Db      1004 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAA 1063
Qy      726 ACCAGCTATGAACTTCATCAGAAATGAACTGATATTATATACCATCAATCTTC 785
Db      1064 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAA 1123
Qy      786 TTTTACCTCAGGACACAGAGAGAAATAGTAAAGATTCAGAGAAATATCTTATAGAAA 845
Db      1124 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAA 1183
Qy      846 AACAAATAGAGAAATCTTATCAACAAATGTTGAAGAGAGAAAGGATTTACATAAGACA 905
Db      1184 AAGGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAGAGAAAA 1243
Qy      906 TCTTGAAGAAAGAAAGATGTTCAATTAACAGAAACAAAGAAAGATTAATCTGCTGA 965
Db      1244 GAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1303
Qy      966 CATACAAATCATATCAATTAAGACAGATTAATATTTTCATGATTTTCAAAATAG 1025
Db      1304 AAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1363
Qy      1026 TAAGTATGAGAGATGAATTAAGTGTGATATGACGATTCATTAATAGTAAAGAGAGA 1085
Db      1364 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAGAAAAAAGAAAA 1423
Qy      1086 TGATGAAGACTTAGACGAATTTAAGCCTATTGTGCAATATGACAAATTTCAAGATGAGA 1145
Db      1424 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAGAAAAAAGAAAA 1483
Qy      1146 AAACATGAGATTTTATTAAGACATGAGAGATTTGATAGAAAAATGAAAAATTTAGTGA 1205
Db      1484 AAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1543
Qy      1206 TTTAGATGAAGGAATGAATAATCATCAGAAATTAATCTGAAGAAAAATTAAGAAAAAG 1265
Db      1544 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAGAAAAAAGAAAA 1603
Qy      1266 AAGAAATATGAATAAACAAGATTAATTTTAAACAAATGATTAAGTTGTATGA 1325
Db      1604 AAAAAAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1663
Qy      1326 TGACATATTTAAAAATATTAAGATTAAGCAGGTTAATTAAGAAAAAGAAAAATTCAT 1385
Db      1664 AAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1723
Qy      1386 AAATCATGTTGTTCTATATTTTGAAGGACATGAAA 1423
Db      1724 AAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1761

```

```

RESULT 11
AJ592058/c 1434 bp DNA linear GSS 15-JAN-2004
LOCUS Arabidopsis thaliana T-DNA flanking sequence, right border, clone
DEFINITION
ACCESSION AJ592058
VERSION AJ592058.1 GI:37941682
KEYWORDS GSS; right border; T-DNA flanking sequence.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

```

```

ACCESSION AJ592058
VERSION AJ592058.1
KEYWORDS GSS; right border; T-DNA flanking sequence.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

```

```

REFERENCE
AUTHORS Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,
1 Chaubin, S., Bechold, N., Cruaud, C., DeRose, R., Pelletier, G.,
Lepliet, L., Caboche, M. and Lecharny, A.
TITLE T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites
JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
PUBMED 12446565
REFERENCE
AUTHORS Balzergue, S.
TITLE Direct Substitution
JOURNAL Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment(s) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
been generated in the framework of the French plant genomics
program 'Genoplante' (http://www.genoplante.com and
http://genoplante-info.infobiogen.fr).
FEATURES
source
1. 1434
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="602D05"
/clone_1b="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Wassilewskija"
1. 1434
/note="T-DNA flanking sequence
right border"
ORIGIN
Query Match 16.0%; Score 237.8; DB 10; Length 1434;
Best Local Similarity 44.5%; Pred. No. 2,5e-29;
Matches 617; Conservative 0; Mismatches 771; Indels 0; Gaps 0;
Qy 5 AACACAAAGCGATCTTAAGACAAAGAGACGTCTTAAGAAAGTTGCAAGAACAA 64
Db 1434 ANANNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1375
Qy 65 GCGATTTAGAACAGTAGCTTGCTTAAGAAAGTTTACAGGACGACAAAGGATTAG 124
Db 1374 NAAAAANNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1315
Qy 125 AACAGAGAGACTTGCTTAAGAAAGTTGCAAGACAAACGATCTTGAACAGAGA 184
Db 1314 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAGAAAAAAGAAAA 1255
Qy 185 GACGTGCTAAAGAAAGTTGCAAGAACAAACAGATTTTGAACAAAGAGACGCTGA 244
Db 1254 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAGAAAAAAGAAAA 1195
Qy 245 AAGAAAAAGTTGCAAGACAAAGCGATTTAAGAACAGTTAGCTTGCTTAAGAAAGT 304
Db 1194 AAAAAANNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1135
Qy 305 TACAGAGACGACGATTTTGAACAAAGAGACGCTTAAGAAAGTTGCAAGAAC 364
Db 1134 NAAAAANNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1075
Qy 365 AACAAAGCGATTTGAACAAAGAGACGCTTAAGAAAGTTGCAAGAACAAAGCG 424
Db 1074 AAAAAANNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1015
Qy 425 ATTTAGAACAAAGAGACCTTGCTTAAGAAAGTTGCAAGAACAAAGCGATTTGAC 484

```

```
Db      1014 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 955
Qy      485 AAGAGAGAGCTGCTTAAGAAAAGTTGCAAGAACAAAGCCATTGAGACAGAGAGC 544
Db      954 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 895
Qy      545 GTGCTTAAGAAAAGTTGCAAGAACAAAGCGATTGAGACAGAGAGAGCTGCTTAAG 604
Db      894 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 835
Qy      605 AAAAGTTGCAAGAGCAGCAAGAGATTGAACTAAAGAGAGGCTGATCGAAAAAAATT 664
Db      834 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 775
Qy      665 TGAAGAAAGAAAAGAAACATGAGATATTTAGCAGAGATTATATGCTGCTTAAGAA 724
Db      774 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 715
Qy      725 TACAGCTATAGAACTTCATCGATGAAAATGAAAGTGATATTATATACCATCAATCTT 784
Db      714 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 655
Qy      785 CTTTACCTCAGAGCAACAGAGGGAATGAGAGATTCCAGGAATATCTATATAGAAA 844
Db      654 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 595
Qy      845 AAACCAATAGAGATCTATTACAACTGTTGAAGAGAGAGGATATACATTAAGAGC 904
Db      594 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 535
Qy      905 ATCTTGAAGAAAAGAAAGATGTTCAATATAACAGAACAAAAGAGATAATCTGCTG 964
Db      534 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 475
Qy      965 ACATACAAATCATACATTAGAGACAGTAATATTTCTGATGTTAATGATTTCAATTA 1024
Db      474 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 415
Qy      1025 GTTATGATGAGATGAATAATGCTGATATGACATTCATTATATGATGAAGAGAG 1084
Db      414 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 355
Qy      1085 ATGATGAGACTTGAAGCAATTTAAGCCTATTTGTCAGATATGCAATTTCCAAGATGA 1144
Db      354 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 295
Qy      1145 AAAACATAGGAATTTATTAAGAACTAGAGATTGATAGAGAAAATGAAAATTTAGATG 1204
Db      294 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 235
Qy      1205 ATTTGATGAGAGATAGAAAATCATCGAGAGATTAATCTGAGAGAAAATTAAGAAA 1264
Db      234 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 175
Qy      1265 GAAAGAAATATGAAAAAACAAGGATATATTTTAAACCAATGATAAAGTTGTATG 1324
Db      174 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 115
Qy      1325 ATGAGCATTTAAAAATATAAAAATGATAGAGGTTTATATAGAAAAGAAAATTTCA 1384
Db      114 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 55
Qy      1385 TAAATCA 1392
Db      54 AATTTAA 47
```

RESULT 12
CL078963 1808 bp DNA linear GSS 31-DEC-2003
LOCUS CL078963
DEFINITION CH216-154E11, RM4.1 CH216 Xenopus tropicalis genomic clone
CH216-154E11, genomic survey sequence.

```
ACCESSION CL078963
VERSION GI:40534876
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1808)
Kremiczki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 17500 Std Error: 0.00
Seq primer: RM4 ctcaaggcatcg9tcgagc
Class: BAC ends
High quality sequence start: 408
High quality sequence stop: 936.
Location/Qualifiers
1..1808
/organism="Xenopus tropicalis"
/mol type="genomic DNA"
/strain="Nigerian frog"
/db xref="taxon:8364"
/clone="CH216-154E11"
/sex="male"
/cell line="Stock 248 F7A2, inbred N7"
/clone lib="CH216"
/note="Vector: pTRABAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN
Query Match 16.0%; Score 237.8; DB 10; Length 1808;
Beet Local Similarity 48.1%; Pred. No. 2.5e-29;
Matches 668; Conservative 0; Mismatches 720; Indels 0; Gaps 0;

Qy      2 AAGAACAAAGCGATTTAGACAGAGAGAGCTGCTTAAGAAAAGTTGCAAGAACAC 61
Db      235 AAGAAANANAAAAAAAAAAAAAAAAAAGNAGAAAGAAAAGAAAAGAAAAGAAA 294
Qy      62 AAGCGATTTGAACAATGACTTGCTTAAGAAAAGTTTCAAGAGCAGAAACGATT 121
Db      295 GAAAAAAAAAAAAAAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 354
Qy      122 TAGAACAGAGAGACTTGCTTAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAG 181
Db      355 AAAAAAAAAAAAAAAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 414
Qy      182 AGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCTG 241
Db      415 AAAAAAAAAAAAAAAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 474
Qy      242 CTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATGACTTGCTTAAGAAA 301
Db      475 AAAAAAAAAAAAAAAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 534
Qy      302 AGTTCAAGAGCAGCAAGCGATTTAGAACAGAGAGAGCTGCTTAAGAAAAGTTGCAAG 361
Db      535 AAAAAAAAAAACAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 594
Qy      362 AACCAACAAGCGATTTAGAACAGAGAGAGCTGCTTAAGAAAAGTTGCAAGAACACA 421
Db      595 AAAAAAAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 654
Qy      422 GCGATTTAGAACAGAGAGAGCTTGCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAG 481
Db      655 AAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 714
Qy      482 AACAGAGAGAGCTGCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAG 541
```

Accession	Source	Position	Sequence	Length
DB	GenBank	715	AAAAAAAAAAAAAAAAAGAAA	774
QY	GenBank	542	GACGTGCTAAAGAAAGTTGCAGAACAAACGATTTAGAACAGAGAGAGCTGCTA	601
DB	GenBank	775	AAAAAAAAAGAAGCG	834
QY	GenBank	602	AAGAAAGTTGCAAGACGACAAAGAGATTGAAACAAGAGAGCTGATCGAAAAA	661
DB	GenBank	835	AAA	894
QY	GenBank	662	ATTGAGAAAGAAAGGAAACATGGAGATATTATGACGAGAGTTTATGTGCTTAA	721
DB	GenBank	895	AAACACAAAGAAAAA	954
QY	GenBank	722	AAATACGCGCTATGAACTTCATCAGAAAAATGAACTGATATTATATACATCAAT	781
DB	GenBank	955	AAA	1014
QY	GenBank	782	CTTCTTTACCTCAGACACACAGAGGAATAGAGATTCCAGAGAAATCTATATAG	841
DB	GenBank	1015	AAA	1074
QY	GenBank	842	AAAAAACAAATAGGAATCTATTCAACAATGTTGAAAGACGAAGGATATACATAAG	901
DB	GenBank	1075	AAA	1134
QY	GenBank	902	GACATCTTGAAGAAAGAAAGATGTTCAATTAACCAAGAACAAAGATTAATCTG	961
DB	GenBank	1135	AAA	1194
QY	GenBank	962	CTGACATACAAATCATCATTAGACACAGTAAATATTCTGATGTTAATGATTTCAA	1021
DB	GenBank	1195	AAAAAAAAAGAAAAAAAAACATAAAAAAAAAAACCAAAAAAAAAAAAAAAAAA	1254
QY	GenBank	1022	TAACTAAGTATGAGAGTGAATTAAGTCTGATATGACGATTCATTATAGTAGAAG	1081
DB	GenBank	1255	AAA	1314
QY	GenBank	1082	AAGATGATGAGACTTAGACGAATTTAAGCCTATTGTCAATATGACAAATTTCCAGATG	1141
DB	GenBank	1315	AAA	1374
QY	GenBank	1142	AAGAAAAATAGGAATTTATTAAGAACTAGAGAGATTGATGAGAAAAATGAAATTTAG	1201
DB	GenBank	1375	AAA	1434
QY	GenBank	1202	ATGATTTGATGAAGGAATAGAAAAATCATCAGAAAGATTATCTGAGAAAAAATTA	1261
DB	GenBank	1435	AAA	1494
QY	GenBank	1262	AAGGAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAAAATGATTAAGTTGT	1321
DB	GenBank	1495	AAA	1554
QY	GenBank	1322	ATGATGACATATTAAATAATTAATAATGATTAAGCAGTTTATAGGAAAAAGAAAT	1381
DB	GenBank	1555	AAA	1614
QY	GenBank	1382	TCATTAATA 1389	
DB	GenBank	1615	AAAAAAAA 1622	

ORGANISM	Xenopus tropicalis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE	1 (bases 1 to 1654)
AUTHORS	Krimitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE	A physical map of the xenopus tropicalis genome
JOURNAL	Unpublished (2003)
COMMENT	Contact: Richard K Wilson Genome Sequencing Center Washington University School of Medicine Email: submissions@watson.wustl.edu Insert Length: 75000 Std Error: 0.00 Seq primer: T7 TAATACGACTCACTATAGG Class: BAC ends High quality sequence start: 468 High quality sequence stop: 812.
FEATURES	Location/Qualifiers
source	1..1654 /organism="Xenopus tropicalis" /mol_type="genomic DNA" /db_xref="taxon:8364" /clone="ISB1-72115" /clone_lib="ISB1" /note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC library Segment 1"
ORIGIN	
Query Match	15.9%; Score 235; DB 10; Length 1654;
Beet Local Similarity	47.6%; Pred. No. 7,3e-29;
Matches	662; Conservative 0; Mismatches 726; Indels 3; Gaps 1;
2	AAGAACAAACAAGCGATCTAGACAGAGACGCTTAAGAAAAGTTGCAAGAACAC 61
29	AAAAACAAAAAAGAAAAACAAAAAGAAAAAGAAAAAGAAACATATAAAAAAGA 88
62	AAAGCGATTTGAACACAGATGACTTGCTTAAGAAAAGTTCAAGAGCAAAAGCGATT 121
89	AAAAACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 148
122	TAGAACAAAGAGAGCTTGCTTAAGAAAAGTTGCAAGAACAAAGCGATCTAGACAG 181
149	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 208
182	AGAGCGTGTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGACGTG 241
209	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 268
242	CTAAAGAAAAGTTGCAAGAACAAAGCG--ATTAGAACAGATGACTTGCTTAAG 298
269	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 328
299	AAAAAGTTCAGAGAGACGAAAGCGATTTAGAACAGAGAGCGTCTAAAGAAAAGTTC 358
329	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 388
359	AAGAACAAACAAGCGATTTAGAACAGAGAGACGTGCTTAAGAAAAGTTGCAAGAACAC 418
389	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 448
419	AAAGCGATTTGAACAGAGAGACTTGCTTAAGAAAAGTTGCAAGAACAAAGCGATT 478
449	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 508
479	TGAAACAAAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAG 538
509	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 568
539	AGAGCGTGTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGACGTG 598
569	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 628

```

QY 599 CTAAAGAAAAGTTGCAGAGCAGCAAAAGATTTTGAACAAGAAAGCTGATACGAAA 658
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 629 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 688
QY 659 AAAATTTTGAAGAAAAGAAACATGAGATATATTACAGAGATTTATATGTCGTT 718
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 689 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 748
QY 719 TGAATAATCCAGCTATAGAACTTCCTCAGAAAATGACGTGATTTATATCCACATC 778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 749 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 808
QY 779 AATCTCTTACCTCAGACAACAGAGGATATAGATTTCAAGAAATATCTATTA 838
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 809 AAAAAAAAAAAAAAAAACCCCCAACAAAAAANNNNNNNNNNNNNNNNNNNNN 868
QY 839 TGAAGAAAACAAATAGAGATCTATTACAACAATGTTGAGAGCGAAGGATATACATA 898
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 869 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 928
QY 899 AAGGACATCTTGAGAAAAGAAAGATGTTCAATTAACCAACAAAAGAGATTAAT 958
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 988
QY 959 CTGCTGACATACAAATCATCATTAAGAGACAGTAATATTTCTGATGTTAATGATTTTC 1018
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 989 AAAAAACAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1048
QY 1019 AAATATAGTATGATGAGATGAATATAGTCTGAATATGACGATTCATTATATGATGAG 1078
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1049 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1108
QY 1079 AAGAGATATGAGCTTGAGCGATTTTGAAGCTATTGCAATATGCAATTTCCAG 1138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1109 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1168
QY 1139 ATGAAGAAACATAGGAATTTTAAAGAACTAGAAATTTGATAGAGAAAATGAATTT 1198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1169 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1228
QY 1199 TAGATGATTTAGTGAAGGATAGAAAATCATCAGAGAAATTTATCTGAAGAAAATTA 1258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1229 AAAAAAATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1288
QY 1259 AAAAAAGAAAGAAATGAAAAAACAAGATTAATTTTAAACCAATGATTAAGTT 1318
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1289 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1348
QY 1319 TGTATGATGACATATTTAAATAATATATGAAGCGTTAATTAAGAAAAGAAA 1378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1349 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1408
QY 1379 AATTCATATAA 1389
    ||| ||| |||
Db 1409 AAAACAGAAA 1419
    ||| ||| |||

```

```

RESULT 14
LOCUS CL066150 1797 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-108B7 Sp6.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CH216-108B7, genomic survey sequence.
VERSION CL066150.1 GI:40522063
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodidae; Xenopus; Silurana.
1 (bases 1 to 1797)
Kremnitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.

```

TITLE
JOURNAL
COMMENT
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 306
High quality sequence stop: 453.
Location/Qualifiers

FEATURES
source
1..1797
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-108B7"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTRABAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN

Query Match 15.8%; Score 234.2; DB 10; Length 1797;
Best Local Similarity 48.9%; Pred. No. 9.8e-29;
Matches 683; Conservative 0; Mismatches 705; Indels 10; Gaps 2;

```

QY 2 AAGAACAAACGATCTAGAACAGAGACGCTGAAGAAAAGTTGCAAGAACAC 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 AAGAAAACAAACCCCATTAATAAAGGNNNNNNNNNNNNNNNNNNNNNN 201
QY 62 AAAGGATTTGAAACAAATAGACTTGTAAAGAAATTTCAAGAGACGAACAGCATT 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 AAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 261
QY 122 TAGAACAGAGAGACTTCTAAAGAAAAGTTGCAAGAACAAACGATCTAGAACAG 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 321
QY 182 AGAGACGCTAAAGAAAAGTTGCAAGAACAAACGATTTAGAACAGAGACGCTG 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 322 AAGAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 381
QY 242 CTAAAGAAAAGTTGCAAGAACAAACGATTTGAAACAAATGACTTGTAAAGAAA 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 441
QY 302 AGTTCAAGAGACGAGCAAGCGATTTAGAACAAAGAGACGCTGAAGAAAAGTTGCAAG 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 501
QY 362 AACCAACAAAGGATTTAGAACAGAGACGCTGAAGAAAAGTTGCAAGAACAAACAA 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 AAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 561
QY 422 GCGATTTGAAACAGAGAGACTTGTAAAGAAAAGTTGCAAGAACAAACGATTTAG 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 562 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 621
QY 482 AACAGAGAGACGCTGAAGAAAAGTTGCAAGAACAAACGATTTAGAACAGAGA 541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 622 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 681
QY 542 GACGTCTAAAGAAAAGTTGCAAGAACAAACGATTTAGAACAGAGACGCTGA 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 682 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 741
QY 602 AAAGAAAGTTGCAAGAGCAAGAGATTTTG-----AACAAAGAAAGCTGATACGA 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 742 AAAAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Qy	65	AAAAAATTAGAAAGAAAAAGAAACAATGAGATTTATTACAGAGATTTATATGTCG	716
Db	802	AAAGAAAAAGAAAAAGAAAAA	861
Qy	717	TTTAGAAATACACGCTATAGAACTTCCATCAGAAATAGACGTGATTTATATACACA	776
Db	862	AAAAAAAAAAGAGAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	921
Qy	777	TCAATCTTCTTACCTCAGACACACAGGGGATATGTAGAGTTCCAGAGAAATATCTAT	836
Db	922	AAAGAAAAAAGAAAAAGAA	981
Qy	837	AATGAAAAAAACAAATAGAGATCTATTACAACAAATTTGAAAGCAGAGGATATATCA	896
Db	982	AAAGAAAAAGAAAAAAGAAAAA	1041
Qy	897	TAAAGACATCTTGAGAAAGAAAGAAAGATGTTTCATATAAACAGAACAAAAAGAAATAT	956
Db	1042	GAAAAAAGAAAAAAGAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGGA	1101
Qy	957	ATCTGCTGACATACAAA-----TCATACATTTAGACAGTAAATATTTCTGATGTTAT	1011
Db	1102	AAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAG	1161
Qy	1012	GATTTTCAATTAAGTAAGTATGAGAGTAAATTAAGTCTGAATATGACATTTCAATTATA	1071
Db	1162	AAAAAGAAAAAAGAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAG	1221
Qy	1072	GATGAGAGAAAGATGATGAGAGACTTAGACAAATTTAAGCTTATGTGCAATATGACAAAT	1131
Db	1222	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA	1281
Qy	1132	TTCCAGATGAGAAACATAGGAATTTATATAGACTAGAGAGATTTGATAGAGAAAAAT	1191
Db	1282	AGAAAAAGAAACGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG	1341
Qy	1192	GAAATTTAGATGATTTAGATGAAGGAATAGAAAAATCATCAGAGGAATTTATCTGAAGAA	1251
Db	1342	AAAAAAGCAAAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1401
Qy	1252	AAAAATAAAAAAGCAAAAGAAATATGAAAAAACAAGGATTAATTTTAAACCAAAATGAT	1311
Db	1402	AACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1461
Qy	1312	AAAAATTGTATGATGACCATATTTAAAAATATATAAAATGATTAAGCAGTTATATAGGAA	1371
Db	1462	AAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1521
Qy	1372	AAGGAAAAATTCATATAA 1389	
Db	1522	AAAAAGAAAAAAGAAAAA 1539	

RESULT	15
LOCUS	CL109409
DEFINITION	ISB1-51H2_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-51H21,
ACCESSION	CL109409
VERSION	CL109409.1
KEYWORDS	GI:40603044
SOURCE	GSS.
ORGANISM	Xenopus tropicalis (western clawed frog)
	Xenopus tropicalis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
	Xenopodinae; Xenopus; Silurana.
REFERENCE	1 (bases 1 to 1829)
AUTHORS	Kremlicki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
TITLE	Mardis,E. and Wilson,R.
JOURNAL	A physical map of the xenopus tropicalis genome
COMMENT	Unpublished (2003)
	Contact: Richard K Wilson
	Genome Sequencing Center

```

FEATURES
source
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: Sp6 ATTAGGTGACACTTAG
Class: BAC ends
High quality sequence start: 856
High quality sequence stop: 936.
location/Qualifiers
1. 1829
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-51H21"
/clone_lib="ISB1"
/note="Vector: pBelobAC11, ISB-1 Xenopus tropicalis BAC
library Segment 1"

```

Query Match	15.8%	Score 234	DB 10	Length 1829
Beet Local Similarity	47.7%	Pred. No. 1.1e-28		
Matches 678	Conservative 0	Matches 742	Indels 0	Gaps 0
3	AGAACCAAAAGCATCTAGAACCAAGAGAGAGCGTCTAAAGAAAAGTTGCACAGAACACA	62		
153	AAAAAGAAAGAGNAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGATAG	212		
63	AAGCATTTAGAACCAAGATAGCTTGTCTAAAGAAAAGTTACAGAGCGCAAGCGATTT	122		
213	AACACCAAAAAGAAAAAAAGCNCAGAAAAAAATATAAAAAAAGAAAAAGAA	272		
123	AGAACAGAGAGACTTCTCTAAAGAAAAGTTGCAGAACCAAGCGCTAGAACAGAA	182		
273	AAAAAAAAAAAAAGAAAGAGAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA	332		
183	GAGACGTCTAAAGAAAAGTTGCAGAACCAAGCGATTTAGAACAGAGAGCGTGC	242		
333	AAAAAAAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAA	392		
243	TAAAGAAAAGTTGCAGAGACCAAGCGATTTGAAACAGATGACTTGTCTAAAGAAA	302		
393	GAAGAAAAAGAAAGAAAAAGAAAAAAAGAAAAAAAGAAAAAGAAAAAGAAAAAGAAA	452		
303	GTTACAGAGAGCAAGCGATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAGAA	362		
453	AAAAAAAAAGACAGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAA	512		
363	ACAACAAAAGCATTTAGAACCAAGAGAGAGCGTCTAAAGAAAAGTTGCAGAACCAAG	422		
513	AAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAGAAA	572		
423	CGATTTGAAACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAACCAAGCGATTTAGA	482		
573	AAAAAAAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAGAAA	632		
483	ACAAGAGAGAGCGTCTAAAGAAAAGTTGCAGAACCAAGCGATTTAGAACAGAGAG	542		
633	AAGAGAGAGAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAGAAAAAGAAAAA	692		
543	ACGTGCTAAAGAAAAGTTGCAGAGAACCAAGCGATTTAGAACAGAGAGAGCGTCTAA	602		
693	AAAAAGAAAAAGAAAGAAAAAGAAAAAAAGAAAAAAAGAAAAAGAAAAAGAAAAA	752		
603	AGAAAGATTGCAGAGAGCAAGAAAGATTTGAAACAAAGAAAGGCTGATACGAAAAAAA	662		
753	AGAAAAAGAAAAAGAGAAAGAAAAAGAAAAAGAAAAAGAGAGAAAAAGAAAAAGAAA	812		
663	TTTAGAAGAAAAAGAACATGAGATATTTAGCAGAGATTTATATGTGCTTTAGA	722		
813	AAAAAAAAAAAAAGAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA	872		
723	AATACGAGCTTAGAATCTTCATAGAAAAATGAACGTGATTTATATCACTCAATC	782		
873	AAAAAAAAAAAAAGAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA	932		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 03:36:00 / Search time 591 Seconds

(without alignments)
937.483 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 1 GAGACACACACACACGATCT.....AATATTTATGAACTATTA 1482

Sequence: 1 GAGACACACACACACGATCT.....AATATTTATGAACTATTA 1482

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA-New:

```
1: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB_seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB_seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB_seq.*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB_seq.*
5: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB_seq.*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB_seq.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
9: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB_seq.*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB_seq.*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77.6	5.2	169495	US-11-121-086-61	Sequence 61, App1
2	74.4	5.0	3181	US-10-793-626-386	Sequence 386, App
3	73	4.9	3543	US-10-793-626-4149	Sequence 4149, App
4	71.6	4.8	171486	US-11-121-086-105	Sequence 105, App
5	69.6	4.7	2655	US-10-793-626-1659	Sequence 1659, App
6	69.6	4.7	3520	US-10-793-626-4160	Sequence 4160, App
7	68.6	4.6	173602	US-11-121-086-25	Sequence 25, App1
8	68.4	4.6	17486	US-11-121-086-105	Sequence 105, App
9	67.6	4.6	3602	US-10-793-626-3582	Sequence 3582, App
10	67.6	4.6	15071	US-10-793-626-2963	Sequence 2963, App
11	67.2	4.5	3840	US-10-793-626-3187	Sequence 3187, App
12	67.2	4.5	4965	US-10-485-517-43	Sequence 43, App1
13	63.2	4.3	1434	US-11-082-746-2	Sequence 2, App1
14	63.2	4.3	6189	US-10-909-125-825	Sequence 825, App
15	63.8	4.2	627	US-10-793-626-693	Sequence 693, App
16	63.8	4.2	627	US-10-793-626-1325	Sequence 1325, App
17	62.8	4.2	2999	US-10-793-626-4439	Sequence 4439, App
18	62.8	4.2	3691	US-10-793-626-3338	Sequence 3338, App
19	62.8	4.2	173601	US-11-121-086-25	Sequence 25, App1
20	60.6	4.1	191091	US-11-121-086-60	Sequence 60, App1
21	59.4	4.0	3305	US-10-793-626-4002	Sequence 4002, App
22	59.4	4.0	150468	US-11-113-908-56	Sequence 56, App1
23	59.4	4.0	193789	US-11-112-908-55	Sequence 55, App1

C	24	59.2	4.0	2796	6	US-10-793-626-4335	Sequence 4335, App
C	25	58.8	4.0	1267	7	US-10-485-517-36	Sequence 36, App1
C	26	58.8	4.0	11896	6	US-11-121-086-84	Sequence 84, App1
C	27	58	3.9	139054	7	US-11-121-086-96	Sequence 96, App1
C	28	57.6	3.9	98862	7	US-11-121-086-76	Sequence 76, App1
C	29	57.6	3.9	182190	7	US-11-121-086-102	Sequence 102, App
C	30	57.4	3.9	207908	7	US-11-112-908-21	Sequence 21, App1
C	31	57.4	3.9	212805	7	US-11-112-908-19	Sequence 19, App1
C	32	57	3.8	151169	7	US-11-121-086-38	Sequence 38, App1
C	33	56.8	3.8	1752	6	US-10-793-626-1357	Sequence 1357, App
C	34	56.8	3.8	3055	6	US-10-793-626-4075	Sequence 4075, App
C	35	56.8	3.8	3819	6	US-10-131-826A-405	Sequence 405, App
C	36	56.6	3.8	3509	6	US-11-077-550-19	Sequence 19, App1
C	37	56.4	3.8	1339	6	US-10-750-185-39902	Sequence 39902, App
C	38	56.4	3.8	120697	7	US-11-121-086-48	Sequence 48, App1
C	39	55.8	3.8	3182	6	US-10-793-626-4047	Sequence 4047, App
C	40	55.6	3.8	747	6	US-10-793-626-463	Sequence 463, App
C	41	55.6	3.8	2036	6	US-10-996-217A-8	Sequence 8, App1
C	42	55.6	3.8	2163	6	US-10-793-626-2057	Sequence 2057, App
C	43	55.6	3.8	3898	6	US-10-793-626-3962	Sequence 3962, App
C	44	55.4	3.7	3068	6	US-10-793-626-3824	Sequence 3824, App
C	45	55.2	3.7	2476	6	US-10-131-826A-489	Sequence 489, App

ALIGNMENTS

```
RESULT 1
US-11-121-086-61
: Sequence 61, Application US/11121086
: Publication No. US20050266459A1
: GENERAL INFORMATION:
: APPLICANT: FOUJSE, TIM S.
: APPLICANT: NIELSEN, KIRSTEN V.
: TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
: FILE REFERENCE: 09138.6000-00000
: CURRENT APPLICATION NUMBER: US/11/121,086
: CURRENT FILING DATE: 2005-05-04
: PRIOR APPLICATION NUMBER: 60/567,570
: PRIOR FILING DATE: 2004-05-04
: NUMBER OF SEQ ID NOS: 107
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 61
: LENGTH: 169495
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: modified base
: LOCATION: (70072)..(70171)
: OTHER INFORMATION: a, c, g, t, unknown or other
: FEATURE:
: NAME/KEY: modified base
: LOCATION: (139457)..(157244)
: OTHER INFORMATION: a, c, g, t, unknown or other
US-11-121-086-61
```

Query Match 5.2%; Score 77.6; DB 7; Length 169495;

Best local similarity 45.1%; Pred. No. 0.0044;

Matches 287; Conservative 0; Mismatches 349; Indels 0; Gaps 0;

Oy	3	AGAACACACACGATCTAGACAGAGACGCTGCTAAAGAAAGTTGCAAGACACACA	62
Db	67358	AC	67417
Oy	63	AAGCATTGTAACACAGATGACTTGTAAAGAAAGTTGCAAGACACACACACACAC	122
Db	67418	GCACATCAAC	67477
Oy	123	AGAACACACAGAGACTTCTAAAGAAAGTTGCAAGACACACACACACACACACAC	182
Db	67478	TCAACACACACAGAGTCAACACACACACACACACACACACACACACACACACAC	67537
Oy	183	GAGACGCTGCTAAAGAAAGTTGCAAGACACACACACACACACACACACACAC	242

```
Db 67538 GATATCATCAAGCATGACGACCAACCAACAGCAAGCAACACGACCACT 67597
Qy 243 TAAAGAAAAGTTGCAAGAACCAAGCATTTAGACAGTCTTGAAGAAA 302
Db 67598 ACCACCAACGAGCAACCACTCTACACAGCAACCAACCAACGCAACCA 67657
Qy 303 GTTACAGAGCAAGAGCATTTAGAACAGAGAGAGTGTAAAGAAAAGTTGCAAGA 362
Db 67658 ACCACCAACCAACAGCAACCAACCAACCAACCAACCAACCAACCAACCA 67717
Qy 363 ACAACAAAGCGATTTAGAACAGAGAGAGTGTAAAGAAAAGTTGCAAGAACAAAG 422
Db 67718 ACTACCAACCAACAGCAACCAACCAACCAACCAACCAACCAACCAACCA 67777
Qy 423 CGATTTAGACAGAGAGAGTGTAAAGAAAAGTTGCAAGAACCAAGCATTTAGA 482
Db 67778 ACAACCAACCAACAGCAACCAACCAACCAACCAACCAACCAACCAACCA 67837
Qy 483 ACAAGAGAGAGAGTGTAAAGAAAAGTTGCAAGAACCAAGCATTTAGACAGAGAG 542
Db 67838 ATTAATAGCAACAGCATTTTCAACACAGCAACCAACCAACCAACCAACCA 67897
Qy 543 ACGTGTAAAGAAAAGTTGCAAGAACCAAGCATTTAGAACAGAGAGAGTGTAA 602
Db 67898 ACCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 67957
Qy 603 AGAAAAGTTGCAAGAGCAACCAAGCATTTAGAACCA 638
Db 67958 GCAGCAATTAACAAAGCAAGCATGTAGAACCA 67993
```

RESULT 2

```
US-10-793-626-3866/c
; Sequence 3866, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3866
; LENGTH: 3181
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3866
```

```
Query Match 5.0%; Score 74.4; DB 6; Length 3181;
Beet Local Similarity 48.7%; Pred. No. 0.012;
Matches 266; Conservative 0; Mismatches 271; Indels 9; Gaps 2;
```

```
Qy 873 TGTGAGAGCAAGGATATACATTAAGAGATCTTGAAGAAAAGATGTTCAAT 932
Db 877 TATTAAAGAAATCCAGATGTACAGAAAGTAATACTAATGCTAATTAAGTACTTCAA 818
Qy 933 AAAACAGAAACAAAAGAAATTAATCTGTCATACAAATCATACATTGAGACAGT 992
Db 817 AACAACTTGTGCAATTTCAATTAAGCTTCAGATGCAAGAGTACGTGAGATTAATA 758
Qy 993 AATATATTCGATGTTAA---TGATTTCAAAATTAAGTAAAGTAAAGATGAATA----- 1044
Db 757 GCAAGAAATTTGATTAATAGCGTGAATTGAAAGAAATTTGCAAGAAATGATCCAGCTTA 698
Qy 1045 -AGTGTGAATATGACGATTTCAATTAATAGATGAAGAAAGATGATGAAGACTTTAGACGA 1103
```

```
Db 697 CATTGAGAAAAAGCGCAAAATTTAGCTTAAACAAATGTAAGAGCAGCAAAATGCA 638
Qy 1104 ATTTAAGCCTATTGTCGAATATGCAATTTCCAGATGAGAAACATAGGAATTTATTA 1163
Db 637 CAAAAAGCTTCAAAAAGATTTAGGAACGTCATTAAGAGAAACAAACACGTAAGC 578
Qy 1164 AGAATAGAAATTTGATAGAAAAATGAATAATTTAGATGATTTAGTGAAGAAATGA 1223
Db 577 TATACAGAAACAGGATTTAAAGACATGAAAAAACTCAAAATATCGAAGAGCAAGTTGG 518
Qy 1224 AAATCATCAGAAATTTATCTGAAGAAAAATTAATAAAGAAAGAAATATGAAAAAAC 1283
Db 517 CTAACTCTCTGTAAATTAATGATCGAAATTTGAGAAAAAGCGCAAACTAGAAAAAGA 458
Qy 1284 AAGGATTAATTAATTTAAACCAATTAATTAAGTTGTATGATGACATTTAAATAATA 1343
Db 457 TATGCTTAAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 398
Qy 1344 TAAAAATGATTAAGCAGTTAATTAAGAAAAAGAAAAATTCATTAATCATTTGTTCAAT 1403
Db 397 TAAAGAAACAAAAACAGCGTGAAGAGCAGAGAAAGCTAGAAAAAAGAAATTTAAAAA 338
Qy 1404 ATTTGA 1409
Db 337 ATATGA 332
```

RESULT 3

```
US-10-793-626-4149/c
; Sequence 4149, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4149
; LENGTH: 3543
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4149
```

```
Query Match 4.9%; Score 73; DB 6; Length 3543;
Beet Local Similarity 43.1%; Pred. No. 0.018;
Matches 581; Conservative 0; Mismatches 750; Indels 18; Gaps 4;
```

```
Qy 11 AAGGATCTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACCAAGCGATT 70
Db 3441 AAAAGAAATGCTAAAAAATGATATTTGATTAAGCAGTGAATGAAGAAATTTCAAA 3382
Qy 71 TAGAACAGATAGACTTCTTAAGAAAAGTTAACAAGACAGCAAGCGATTTAGAACAG 130
Db 3381 GAATTAATGAATGCTACATATCAATTAAGAAAGCAAAATCGTAAATTTGAAGAGCTA 3322
Qy 131 AGAGACTTGTAAAGAAAAGTTGCAAGAACCAAGCGATCTAGAACAGAGAGAGCGTG 190
Db 3321 AGATTGAAGCAAAAGATTAATTTCA---ACGCAATGTACTAGAGATCAAGTAAATGA 3267
Qy 191 CTAAAGAAAAGTTGCAAGAACCAAGCGATTTAGAACAGAGAGAGCGTGTAAAGAA 250
Db 3266 AGCGAAAATTAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3207
Qy 251 AGTTCAAGAACCAAGCGATTTAGAACAGATAGACTTGTCTTAAGAAAAGTTAGCAAG 310
Db 3206 TGAAGCTAGACAGCAGTACGAATTAAGCAA-ATGAACAGATTAATCATATTTCAAAACA 3148
```



```
Db      2476  AAAAAACAAGACACATCTTAAAAATGAACATGAAGAGTTGAATTTGAAAAATGA 2535
Qy      780  ATCTCTTTTACCTCAGACACAAGAGGAAATAGTAGAATTCAGAGAAATATCTATAT 839
Db      2536  TGGATATCAAGATGTAAGAGTAAGAAACCTTAAAAAGAAAAACAAATCATTTAATGA 2595
Qy      840  AGAAAAACAATAGAGATCTATTACACAAATGTTGAAGACGAAGGATATACATTA 899
Db      2596  GATACAAACAATTGAAGCAACTAGAAAGATATTGAAAGATATCACATATATCAAA 2655
Qy      900  AGGACATCTGGAAGAAAAGAAAGATGCTCAATTAACCGAAGAAAAGAGTAATATC 959
Db      2656  AGAAGAAAAGCTTCGACACTCAACACAACAACACATCATCAAAAACATCTGATTT 2715
Qy      960  TGCTGACATACAAATATATCATTTAGAGACAGTAATATTTCTGATGTTAATGATTTCA 1019
Db      2716  AGCTGTTGTTAAGAGGAAATTAATCGCAAAAGCAAGTTTATGAGCTTTAGATTAACA 2775
Qy      1020  AATAGTAAGTATGAGATGAATTAAGTCTGAAATATGACGATTCATTAATAGTGAAGA 1079
Db      2776  ACTTAGGATTCAGAACGTCAAAAAATTTGAAGTAAGTAAGTAAGTAAGTAAGTTC 2835
Qy      1080  AGAAGATGATGAAACCTTAGCGAATTTAAGCCTATTTGCAATATGACAATTCGAGA 1139
Db      2836  AGATGAATATGATGGTAAAGATGCTTTTGAAGAGTTGAAGAGCAAAATTCGACACAGA 2895
Qy      1140  TGAAGAAAACATAGGAATTTTAAAGAACTAGAAAGATTTGATAGAGAAAATGAAAAATTT 1199
Db      2896  AAAATGTAAGCAAAATTTAATCAACAATTAAGTGAATTTAAAGCAAGCTAAAGATCT 2955
Qy      1200  AGATGATTTAGATGAAGATAGAAAAATCATCAAGAAATTTATCTGAAGAAAAATPAA 1259
Db      2956  TAAATGAAGAAATCGAAATTAATGAAGTCAGAGCTTCAAAAATGTCATCAAGATATACCTTC 3015
Qy      1260  AAAAGGAAGAAATATGAAAAACAAAGATATATTTTAAACCAATG 1309
Db      3016  TATAGAAATCATATATCAAGATATTTAAAGCAAAACATCAAGCTAGATG 3065
```

RESULT 7

```
US-11-121-086-25/c
; Sequence 25, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121, 086
; PRIOR FILING DATE: 2005-05-04
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25
; LENGTH: 173602
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-25
```

```
Query Match      4.6%; Score 68.6; DB 7; Length 173602;
Best Local Similarity 44.5%; Pred. No. 0.07;
Matches 365; Conservative 0; Mismatches 449; Indels 7; Gaps 2;
```

```
Qy      656  AAAAAAATTTGAAGAAAAAGAAACATGAGATATTTTTCGAGAGATTTATATATGCTC 715
Db      13631  AATAATATATTAACATATTAATAATAATAATAATAATAATAATAATAATAATAATA 136772
Qy      716  GTTGAAGAAATACAGCTATAGAACTTCATCAGAAAGATGAGATATATATATACAC 775
Db      136771  TAAATATATTAACATATTTATATATATATATATATATATATATATATATATATAT 136712
```

```
Qy      776  ATCAATCTTCTTACTCAGACACAGAGGAATAGTAGATTCAGAAATATCTA 835
Db      136711  ATATAAAAATATAT-----TATATATATATATATATATATATATATATATATAT 136657
Qy      836  TAATGAAAAAACAATAGAGATCTATTACAACAATGTTGAAGACGAAGGATATATAC 895
Db      136656  ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 136599
Qy      896  ATAAAGACATCTTGAAGAAAAGAAAGATGCTCAATTAACCAAGCAACAAAAAGAGATA 955
Db      136598  ATAAATATATATATATATATATATATATATATATATATATATATATATATATAT 136539
Qy      956  AATCTGCTGACATACAAAATCATACATTTAGAGACAGTAATATTTCTGATTAAGATT 1015
Db      136538  ACTATATAAATATATATATATATATATATATATATATATATATATATATATATAT 136479
Qy      1016  TTCAAATAGTAAGTATGAGATGAAATTAAGTCTGAAATATGACATTCATTAATGATG 1075
Db      136478  TATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 136419
Qy      1076  AAGAAGAAATGATGAAGACTTGAAGCAATTTAAGCCTATTTGCAATATGACAATTTCC 1135
Db      136418  TTTTATATATATATATATATATATATATATATATATATATATATATATATATAT 136359
Qy      1136  AAGATGAAGAAACATAGCAATTTATTAAGACATGAGAGATTTGATGAGAAAAATGAAA 1195
Db      136358  TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 136299
Qy      1196  ATTATGATGATTTAGATGAAGAAATGAAAAATCATCGAAGAAATTTCTGAAGAAAAA 1255
Db      136298  TATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 136239
Qy      1256  TAAAAAAGAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATATATA 1315
Db      136238  TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 136179
Qy      1316  GTTTGATGATGAGCAATTTAAAAATATATAAATGATAGCAAGTTATATAGGAAAAAG 1375
Db      136178  TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 136119
Qy      1376  AAAAAATCATTAATCATCTGTTTCATATATTTGACGAGACAAATGAAATTTTACAGATC 1435
Db      136118  ATATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 136059
Qy      1436  TGAATGATATCTGAAGATATTAACATTAATATTTTATGAAA 1476
Db      136058  AATATATGTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 136018
```

RESULT 8

```
US-11-121-086-105
; Sequence 105, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121, 086
; PRIOR FILING DATE: 2005-05-04
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 105
; LENGTH: 171486
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-105
```

```
Query Match      4.6%; Score 68.4; DB 7; Length 171486;
Best Local Similarity 46.2%; Pred. No. 0.075;
Matches 370; Conservative 0; Mismatches 421; Indels 9; Gaps 4;
```

[illegible]

```

; LENGTH: 3602
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3582

```

Query Match	4.6%	Score 67.6;	DB 6;	Length 3602;
Best Local Similarity	43.1%	Pred. No. 0.095;		
Matches 438; Conservative	0;	Mismatches 569;	Indels 9;	Gaps 2.

QY	333	AGAGAGCGTGTCTAAAGAAAGTTGGCAAGAACAAAGAGATTTAGAACAGAGAGACG	392
Db	2208	AGTAAATCCGTGTAAACCAAGTTTAAATTTATAGAAAGCGTGATTTATTAATTTAAACGTC	2149
QY	393	TGCTAAAGAAAGTTGGCAAGAACAAAGCGATTTTGAAACAAGAGACTGCTTAAGA	452
Db	2148	ATTTAACAAAGCTTTAAAGAGCTAAGGCGTATTAATTAATAATGCGGTACAAATGT	2089
QY	453	AAAGTGGCAAGAACAAAGCGATTTTGAAACAAGAGAGCGTGTAAAGAAAGTTGCA	512
Db	2088	CAATATCAATGACATTCACATCTTTTAAACAAATATGATATGTCTTAAGCCAAATTAA	2029
QY	513	AGAAACAACAAAGCATTTAGAACAGAGAGCGTCTTAAAGAAAGTTGCAAGAACACA	572
Db	2028	TGCTGAACGACGTCTAAAGAACATCAACAAATCTGAAGTATTTATTTAAAGATTT	1969
QY	573	AAGCATTTTGAACAAGAGAGACGTGTAAAGAAAGTTGCAAGAGAGCAAGAGATTT	632
Db	1968	AGATATACCTTAATATATGCTCAAAAAGCTGCATTA-----TTAATACGATTTAGCGCTC	1915
QY	633	AGAACAAAGAAAGCGTGTATCGAAGAAAGAAATTTGAAAGAAAGAAAGCAATGAGATAT	692
Db	1914	TAAAGACATTTAAATTAATTAATCAATCGTTGATTAATGCAATAGAAATTAATATGTCAT	1855
QY	693	ATTACGAGAGATTTATATGTGCTGTTTGAATATCCGCTATAGAACTTCATAGAAA	752
Db	1854	GCAAGGTTTAAAGAACATGTAGCTCAATTAACAGCAATCAACAAACAAATTTGATATA	1795
QY	753	TGAACGTGATTTATATATACACATCAATCTCTTACCTCAGGACCAAGAGGAATAG	812
Db	1794	TTTAATGTGATGAGACCTTAAATATCAATATGATTAACGCTATCAACTTTAGGAAATTA	1735
QY	813	TAGAGATTCGAAGAAATATCTATATATAGAAAGAAACAAATAGAAATCTATTAACAACA	872
Db	1734	TGTTCTTGACAAAGAAACGCTAACAAATPAAAGACGTAATATCAATATGGAATGATTTCA	1675
QY	873	TGTTGAMGACGAAGGATATATCATTAAGACATCTTGAGAGAAAGAAAGATGTTCAAT	932
Db	1674	AAACATGATGATCTAGAGACCTTCTAATATGAAATGGAAGCTTAAAGATGCTCAAC	1615
QY	933	AAACCGAACAAGAAAGATTAATCTGCTGACATCAAAAATCATTCATTAGAGACAGT	992
Db	1614	AAAGGACATATATGACATTAATAGATACGCTCAACGCTCACTTATGAAATTTGAACAGC	1555
QY	993	AAATTTCTGATGTTATATGATTTTCAAAATAGTAGATGAGAGTGAATATAGTCTGA	1052
Db	1554	TAAATCAACATCAAAATCTTAAGCTCAAGCAAAACAATGTTAATATGAGAA--GCTAG	1498
QY	1053	ATATGACGATTCATTAATATAGATGAAGAAAGATGATGAGACCTTAGACGAATTTAAGCC	1112
Db	1497	AAAAGCGTTTCTATATATTAATATCAACGCAACATCAAAATGATTTAGTTATCAAGCAAGA	1438
QY	1113	TATTTGGAATATGACAAATTCGAAGATGGAAGAAACATATGGAATTTATTAAGAATAGA	1172
Db	1437	TGAAGGGAATCTGCAATTTGAACACATACGACGATGAATTTACTTAAGCAAACTTGA	1378
QY	1173	AGATTTGATAGAGAAATGAAATTTTATAGATTTTATGATGATGAGAAATAGAAAATCATC	1232
Db	1377	TGCTATCAAAATGATTTGACAAAAGTTGAAGATATTAATCATCTTAATTTATGCAAAATCC	1318
QY	1233	AGAGAAATTAATCTGAAGAAAGAAATTAAGAAAGAAAGAAATATGAAAAACAAAGATTA	1292

Db 1317 AAATTGATCAATGAAAGAAAAAAATTAACCTAATATCTCAATATTAATAGTTAGTAAATG 1258
 Oy 1293 TAATTTTAAACCAATGATAAAAGTTTGTATGATGAGCATATTTAAAAATATPAAA 1348
 Db 1257 AATTAAAGATGAATTCACACAGCTAATAACAAACAATAAGAAAAAGCTACAA 1202
 RESULT 10
 US-10-793-626-2963
 : Sequence 2963, Application US/10793626
 : Publication NO. US20050255478A1
 : GENERAL INFORMATION:
 : APPLICANT: KIMMERLY, WILLIAM JOHN
 : TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 : FILE REFERENCE: PUS480US
 : CURRENT APPLICATION NUMBER: US/10/793,626
 : CURRENT FILING DATE: 2004-03-04
 : PRIOR APPLICATION NUMBER: 60/164,258
 : PRIOR FILING DATE: 1999-11-09
 : NUMBER OF SEQ ID NOS: 4472
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 2963
 : LENGTH: 15071
 : TYPE: DNA
 : ORGANISM: Artificial Sequence
 : FEATURE:
 : OTHER INFORMATION: Description of Artificial Sequence: synthetic
 : OTHER INFORMATION: nucleic acid sequence
 US-10-793-626-2963

Query Match	4.6%	Score	67.6	DB	6	Length	15071
Best Local Similarity	43.1%	Pred. No.	0.035				
Matches 48		Conservative	0	Mismatches	569	Indels	9
						Gaps	2

333 GGAGAGACGCTCTAAAGAAAAGTTCGCAAGAACCAACAAAGCGATTGTGAACAAGAGAGACG 392

12864 AGTAATCTCGTGTAAACCAAGTTTAAATTTATAGAAACGCTGATTATTAATTTAAACGTCA 12923

393 TGCTAAAGAAAAGTTGCAAGAACCAACAAACGCAATTGTGAACAAGAGACCTGCTTAAGA 452

12924 ATTTAACCAAGCTTTAAAGAAAGCTTAAGCGCTTTTAATTAATAATATGCGGTCAAAATGT 12983

453 AAAGTTCGAGAACCAACAAACGCAATTGTGAACAAGAGACGCTCTAAAGAAAAGTTGCA 512

12984 CAATATCAATACATTCACACATCTTTTAACCAAAATGATATATGCTTAAGACCAATTAAA 13043

513 AGAACCAACCAAGCAATTGTGAACAAGAGACGCTCTAAAGAAAAGTTGCAACACACA 572

13044 TGTGTGAACGAGCTCTAAAGAACATCAACAAAATCTGAAGTATTTATTATTAAGAAATT 13103

573 AAGCGATTTAGACACAGAGACGCTGCTTAAGAAAAGTTGCACAGACGACCAAGAGATT 632

13104 AGATTACTTATATATGCTCAAAAAGCTGCATTA-----TTAATCAGATTAGACGTC 13157

633 AGAACCAAGGAAGCTGATACGAAAAAAAATTTGAAGAAAAAGAACATGAGATATC 692

13158 TAAAGACATTTAAATTAATTAATCAAAATCGTTGATTAAGCAATGAAATTAATATANTCTAT 13217

693 ATTGACAGAGATTTATATGTCGTTTGAATAATCCACGATATGAACCTTCATCAGAAAA 752

13218 GCAGAGTTTAAAGAACATGTAGCTCAATTACACGCACTACAAAGACAAACATTGAAT 13277

753 TGAACGTGATATTATATACACATCAATCTTCTTACCTCAGAGACACAGAGGAATAG 812

13278 TTTAATCTCTATGAAGACCTTTAAATATCAATATGATTAGCGTATCAACTTACGAATTA 13337

813 TAGAGATTCCAGGAAATATCTATATATAGAAAAACAATATAGAAATCTATTCAACAA 872

13338 TGTCTTGTACAAAGAAAACGGTACAAATATAAAGACGTATATCTATATGGAATGATTCA 13397

873 TGTGTAAGACGAGGATATATCAATTAAGACATCTTGAAAGAAAAAGATGCTTCAAT 932

Db	13398	AAACATGATGATGCTAGAGACACTCTTAATGAAATTGAAAGACTTAAAGATGCTCAAC	13457
Oy	933	AAAACCAAGAAAAAGAAATTAATCTGCTGACATCAAAATTCATACCTTTAGACAGT	992
Db	13458	AAAAGCATAAATGACATTTAAAGATCGCTCAACCGCTCAACTTGATGAATGAACAGC	13517
Oy	993	AAATATTTCTGATGTTAATGATTTTCAAATAAGTAAGTATGAGATGGAATAAGTGCTGA	1052
Db	13518	TAAAGCAACATCAAAATTTCTAAAGCTCAACCAAAACAAATGATGAATGAGAA--GCTAG	13574
Oy	1053	ATATGACGATTCATTAATAGATGAAGAAGATGATGGAAGCTTAGAAGATTTAAGCC	1112
Db	13575	AAAAGCGTTTCTAAATTAATTAATCAGCAACATCAAAATGATTTAGTTAACTAAGCAAAAGA	13634
Oy	1113	TATTTGCAATATGACATTTCCAAAGTGAAGAAATCAGGAATTTATPAAGAAGTGA	1172
Db	13635	TGAAGGCAATCTGCATATGAACACATACATGCAATGAATTAACCTAAAGCAAACTAGA	13694
Oy	1173	AGATTGATGAGAAAAATGAATAATTAGATTTAGATGGAAGATTAAGAAAAATCATC	1232
Db	13695	TGCTAATCAATATGTTGACCAAAAGTTGAAGATTAATAATCACTTAATTAATGTCAAATCC	13754
Oy	1233	AGAAGAATTAATCGAAGAAAAATTAATAAAAGAAAGAAATATGAAAAAACAAGATPA	1292
Db	13755	AAACTTATCAAAATGAGAAAAATTAATCTAAATATCTCAAAATTAATTAATTTAGTAAATGG	13814
Oy	1293	TAAATTTAAACCAATGATAAAGTTTGTATGATGACATATTTAAAAATATPAAA	1348
Db	13815	AATTAAGATGAATTCACACACTTAACCAACACAAATAGAAAAATGCTACAA	13870

RESULT 11
US-10-793

US-10-793-626-3187
; Sequence 3187, Application US/10793626

```

? GENERAL INFORMATION:
? APPLICANT: KIMMERLY, WILLIAM JOHN
? TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
? FILE REFERENCE: PUJ480US
? CURRENT APPLICATION NUMBER: US/10/793,626
? CURRENT FILING DATE: 2004-03-04
? PRIOR APPLICATION NUMBER: 60/164,258
? PRIOR FILING DATE: 1999-11-09
? NUMBER OF SEQ ID NOS: 4472
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 3187
? LENGTH: 3840
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: synthetic
? US-10-793-626-3187

```

Query Match 4.5%; Score 67.2; DB 6; Length 3840;

```

Best Local Similarity  45.7%;  Freq. NO. 0.11;
Matches  425;  Conservative  0;  Mismatches  543;  Indels  12;  Gaps  2;

```

QY AACCAAGAGAGACGCTCTAAAGAAAAAGTTGCAGAAACAACAAAGCATTTAGAACAAAGAGA 439

QY 380 AACCAAGAGAGACGCTCTAAAGAAAAAGTTGCAGAAACAACAAAGCATTTAGAACAAAGAGA 439

Db 137 AAGCCAAATTAATATGCGAGACATCAACCCAGGTGTCAAAACCTATCAAAGCAGACGCA 196

QY 440 GACTTGCTAAAGAAAAAGTTGCAGAACAAACGAGATTTAGAACAAAGAGACGCTGCTTA 499

Db 197 TAACCTCTTTATCTCGAATTATATGCACAAGTTGTTGAGAAAGAGCTCGCAAGAAATGCA 256

QY 500 AAGAAAAAGTTGCAGAAACAACAAAGCATTTAGAACAAAGAGACGCTCTAAAGAAAAGT 559

Db 257 TCGAACAAAGGCAACACAAACAAAGCAATTTATTAATATA--ATGATTAATGCTACAG 313

QY 560 TCCAGAAACAACAAAGCGATTTAGAACAAAGAGACGCTCTAAACAAAGTTGCAAGAGC 619

Db 314 ATGAAAGAAAGAGGTCGCCAACAAATTTAGTTATGCTCTCAAAAACAAAAATCATTTAGATA 373

```
QY 620 ACCAAGAGATTAGAACAAAGGAGCTGATACGAAAAAATTTAGAAAGAAAAAG 679
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 AATATTACTCTTATCTTCAAAATTAATGATGTTGAAAAATCTTAAAGTAGAGAAATTAATG 433
QY 680 AACATGAGATATATTAGCAGAGATTATATGTCGTTTAAATAATCAGCTATAGAAC 739
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 434 AATATGCTTAACGTTTATCCAGAACCGCTGTTTAAGTCAAAACCAAAAAAGATATTTGATC 493
QY 740 TTCCTACGAAAAATGAAAGTGATTTATATACCATCAATCTTCTTAACTCGAGGACA 799
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 494 AAAAATCTGCGACACGATTAAATCAAATTCAAACGATCAAACTGCTACATCGAGGAAA 553
QY 800 ACAGAGGATATAGTAGAGATTCCAAGGAATATCTTAAATAGAAAAACAATAGAGAT 859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 554 AAGAGCGGCTATTCATTTGGCAAAATCAAAAATCAATAGAACAGAAACAGCAATTCAAA 613
QY 860 CTATTACAACAATGTTGAAGACGAAAGGATATACATAAGGACATCTTGA----- 911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 614 ATGAACATATGTAACATGGTGTGCGCAAGCTAAATCTAACGCAATTCATGAATTTGAT 673
QY 912 -AGAAAGAAAGATGGTTCAATTAACCGAACAACAAAAGATAATCTGCTGACATAC 970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 674 TGGTATGCGCAGATGCGCACAAAAATCTGATGCTAAACAAAGTATCGATAATTAATATA 733
QY 971 AAAATCATATCATTAGACAGTAAATATTTCTGATGTTATGATTTTCAATATAGTAAGT 1030
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 734 ATGAGCAAACTAACTATCAACACTACACAGATCAACAGATGAAGAAAAGCAAAAAG 793
QY 1031 ATGAGATGAATTAAGTCTGTAATATGACGATTCATTAATATGATGAAGAAAGATGATG 1090
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 794 CATTGATTAATTAATAAATAGCTAAAGATGACAGATACAAACAAAGTTGTCAGCGCAA 853
QY 1091 AAGACTTAACGAATTTAAGCCATTTGTGCATATGACAAATTCACAAAGTGAAGAAAACA 1150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 854 CAACCAACAAAGATCTGATGCAAAAACAGAGCTATAGATACGATTAATATTTCAAG 913
QY 1151 TGAATTTATTAAGAACTAGAAAGATTTGATAGAGAAAAATGAAATTTAGATGATTTAG 1210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 914 CAAATGTTCAAAAAAACCATCCGCTCGAGTGAATTAATTCMAAGTTTGAAGATTTAA 973
QY 1211 ATGAAGAAATAGAAAAATCATCAGAAATTAATCTGAAGAAAAAATTAAAAAAGAAAGA 1270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 974 ACGCTCAATCAATGACACGCCCAATGCTACAGAAAGAAAAACAAGATGCAATTCAAA 1033
QY 1271 AATATGAAAAACAAGATTAATTAATTTAAACCAATATGATAAAGTTGTATGATGAC 1330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1034 GATGATGTAATAAAGATGAAGTTAAGATCTTAATTAATCAAGATAGACGTGCAATG 1093
QY 1331 ATATTAAAAAATTAATAAAT 1350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1094 AAGTTGACACGACAAAAT 1113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-10-485-517-43/c
; Sequence 43, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexis Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P1062360
; CURRENT APPLICATION NUMBER: US/10/485,517
; PRIORITY FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825,9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349,9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: Patentin version 3.1
```

```
; SEQ ID NO 43
; LENGTH: 4965
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-485-517-43
Query Match 4.5%; Score 67.2; DB 6; Length 4965;
Best Local Similarity 45.1%; Pred. No. 0.11;
Matches 341; Conservative 0; Mismatches 403; Indels 12; Gaps 2;

QY 563 AAGAACACAAAGCGATTTAACAAGAGACGCTGAAGAAAAAGTTGACAGCGAC 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3843 AAGAGAAAAAGCAGCAAAAAGATTGTTTAAAGCTTAAAGAAAAAGCGTTACAGATA 3784
QY 623 AAAGATTTTGAACAAAGAAAGCGCTGATACGAAAAAATTTAGAAAAAAGAAAGAAC 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3783 TCTTAAATGCAACAAACATATGATGTTCGCAAAATTAAGATCAACAGTGTGCTGATA 3724
QY 683 ATGAGATATATTAGCAGAGATTTATATGTCGTTTGAATAATACAGCTATAGACTTC 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3723 TTCAAGTATTAACGACATACAAACATTAAGAATGTTGCAAAAGATGATTAAGCAACA 3664
QY 743 CATCAAAAAAGAGTGAATTTATATACCAATCAATCTTCTTACCTCAGACACACA 802
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3663 AAGCAACGAAACAAAACGCTTATTTGCAAACTGCGATGCGACTAGAAAGAAAAG 3604
QY 803 GAGGGAATAGTAGAGATTCCAAGAAATATCTATATAGAAAAACAATAGAGAACTGA 862
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3603 AACCAAGCAATTCACAACTAGACGCAATTTAACCAAGTATCAAAATATTTGAATAAG 3544
QY 863 TTACAACAATATGTTGAAGAGAGAGGATATACATAAAGACATCTTGAAGAAAAAGAA 922
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3543 CACAGTCAATCGATGATTAACAACCTGCAAAAAGATATGCAATTCACAAATTCGCCAA 3484
QY 923 ATGCTTCAATTAACCAACAAACAAAAGAAATTAATCTG-----CTGACATACAA 973
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3483 TTCAAGCATCAACAGATGTTAAACGAATGCAAGCGGAAATGCTAACTGAATGCAAA 3424
QY 974 ATCATCATTTAGAGACAGTAAATATTTCTGATGTTATGATTTTCAATTAAGTAAGTATG 1033
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3423 ATTAATTAATCTGAATCTTAATTAATATGAGACTATCAATGAAGAAAAAGTAAACGATA 3364
QY 1034 AGGATGAATATAGTCTGAATATGACGATTCATTAATATGATGAGAGAAAGATGTAAG 1093
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3363 TTGACACAGTTAGACACATATGAAAGAGTTTAAATTAATTAATGACAGAACTACTA 3304
QY 1094 ACTTGAAGAAATTTAAGCTTATGTGCAATATGACAAATTTCCAGATGAAGAAAACATAG 1153
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3303 CAGGTGATGTACTACTGCTAAAGATACAGACAGTACAAAAGTTCAACAACTTCACAA 3244
QY 1154 GAATTTATTAAGAACTAGAAAGTGTGATAGAGAAAAATGAAAAATTTAGATGATTTAGATG 1213
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3243 ATCTGTTTAAGAAACACAGAGTAAATAAAGATTTAGATCAAGCTGACGTGATTAAGAAA 3184
QY 1214 AAGGAATGAATAATCATCAGAAAGATTAATCTGAAGAAAAAATTAAAAAAGAAAGAAAT 1273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3183 CACAAATAGAACAAACACCAATG---CATCAACACAGAAATTAATGATGCAAAAACAAG 3127
QY 1274 ATGAAAAACAAGATTAATTTTAAACCAAAAG 1309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3126 AAGTTGATCTGAATTAATCAAGCAAAAACAATG 3091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-11-082-746-2/c
; Sequence 2, Application US/11082746
; Publication No. US2005024461A1
; GENERAL INFORMATION:
; APPLICANT: Sundstrom, Paula
; TITLE OF INVENTION: NUCLEIC ACIDS ESSENTIAL FOR EXPRESSION OF HYPRAL-SPECIFIC GENES
; FILE REFERENCE: DC-0300
; CURRENT APPLICATION NUMBER: US/11/082,746
```

```
; CURRENT FILING DATE: 2005-03-17
; PRIOR APPLICATION NUMBER: US 09/725,010
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 60/167,672
; PRIOR FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Candida albicans
US-11-082-746-2

Query Match      4.3%; Score 63.2; DB 9; Length 1434;
Best Local Similarity 52.6%; Pred. No. 0.37;
Matches 160; Conservative 0; Mismatches 143; Indels 1; Gaps 1;

QY 1107 TAAAGCTATTGTGCAATATGACAAATTTCCAAAGATGAGAAAACATAGAAATTTATAAGA 1166
Db 305 TTAACCACTGAATTAAGCTTGAACCTAATCTTTGTAAGAAAAGGGGATTTTATTAAAA 246
QY 1167 ACTGAAGATTGTAGAGAAAATGAAAATTTAGATGATTAGATGAGAAATAGAAA 1226
Db 245 GCTATCAACTATGAAACGAAAACAGAAAACAAAGCAATGAGAGAAAGAAAAGATTA 186
QY 1227 ATCTTCGAAGAATTAATCTGAGAAAATTAATAAGAAAAGAAAAGAAATATGAAAAACAA 1286
Db 185 ACAATTTGAACAAACAAATTTTAAACCTAAAGATCTATCACTTTAATGTAGTAA 126
QY 1287 GGATATATATTTTAAACCAATGATTAAGTTGT-ATGATGACATATTTAAAAATATA 1345
Db 125 CAAACCAATATTTCAATTAATGAGCTCTTGCAATTAATATTGAAACCAGAAATTTGAAA 66
QY 1346 AAAATGATTAAGCAGTTAATTAAGAAAAGAAAATTCATTAATTCATTTGTCATATAT 1405
Db 65 ACATCAAAATTAAGAAAAAAGAAATCCCAAAAAAAGAAACAAACCTTAGATCAGTTAT 6
QY 1406 TTGA 1409
Db 5 TAGA 2

RESULT 14
US-10-909-125-825
; Sequence 825, Application US/10909125
; Publication No. US20050261218A1
; GENERAL INFORMATION:
; APPLICANT: Esau, Christine
; APPLICANT: Benneit, C. Frank
; APPLICANT: Freiler, Susan M.
; APPLICANT: Grifley, Richard H.
; APPLICANT: Baker, Brenda F.
; APPLICANT: Vickers, Timothy
; APPLICANT: Marcussen, Eric G.
; APPLICANT: Koiler, Eric
; APPLICANT: Swayze, Eric
; APPLICANT: Jain, Ravi
; APPLICANT: Bhat, Balkrishen
; APPLICANT: Peralta, Eileen
; TITLE OR INVENTION: Oligomeric Compounds And Compositions For Use In Modulation
; FILE REFERENCE: 1815080-100 (CORE0016US)
; CURRENT APPLICATION NUMBER: US/10/909,125
; CURRENT FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: US 60/492,056
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US 60/516,303
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: US 60/531,596
; PRIOR FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: US 60/562,417
; PRIOR FILING DATE: 2004-04-14
```

```
; NUMBER OF SEQ ID NOS: 2184
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 825
; LENGTH: 6189
; TYPE: DNA
; ORGANISM: H. sapiens
US-10-909-125-825

Query Match      4.3%; Score 63.2; DB 6; Length 6189;
Best Local Similarity 42.2%; Pred. No. 0.37;
Matches 517; Conservative 0; Mismatches 688; Indels 21; Gaps 2;

QY 133 AGACTGCTAAAGAAAAGTTGCAAGAACAAACAGCATCTAGAACAAAGAGACGCTGCT 192
Db 1894 ACATTGAAAATTTACAGTTGTCACACCTGAAAAAATGATTAAGAGATTAACATTTAAA 1953
QY 193 AAAGAAAAGTTGCAAGAACAAACAGATTTTGAACAAAGAGACGCTGCTAAAGAAAAG 252
Db 1954 AACATTAATTTGAAGCTAAAGAGAGAAAAGAAAAGCATTAAGAAAATTTGAAGGT 2013
QY 253 TTGCAAGAACAAACAGCATTTTGAACAAAGATGACTTGAAGAAAAGTTACAAAG 312
Db 2014 GAAAGGAAAAATACAAACTAAGATAGTCCAAAGAACTGCAAGAGAGTGTGAATTT 2073
QY 313 CAGCAAGCATTTTGAACAAAGAGACGCTTAAGAAAAGTTGCAAGAACAAACAG 372
Db 2074 GATAGCAATTTTGAAGAGAAATTTTAAAGTATGATAAAGTGAAGATCTTTTAA 2133
QY 373 GATTTAGAACAAAGAGACGCTTAAGAAAAGTTGCAAGAACAAACAGATTTAGAA 432
Db 2134 AATATGACATGATATCTTAACATTTAGAAAATTAACAAATTTGAAAATTAACATCAA 2193
QY 433 CAAAGAGACTTGCTTAAGAAAAGTTGCAAGAACAAACAGATTTTGAACAAAGAGAG 492
Db 2194 GATGATTAATCAACCAAGGAAAAGCATGTCTCAAGAGAGGAACTTTAAAGGAGAG 2253
QY 493 CGTCTAAAGAAAAGTTGCAAGAACAAACAGATTTTGAACAAAGAGAGACGCTTAA 552
Db 2254 GACAGATTAAGAAAAGGAAAGGAGAAATCTTTTAAAGGAGAGAAAATTAAGATCTTAA 2313
QY 553 GAAAAGTTGCAAGAACAAACAGATTTTGAACAAAGAGAGACGCTTAAAGAAAAGTTG 612
Db 2314 GAAAGAGAGAAAATCAATCCACAGATTAAGATCAAGATTTCTTTGGGATAGAT 2373
QY 613 CAAAGACGCA-----AAGATTGAACAAAGGAGGCTGATCGAAAAAAT 663
Db 2374 GCCATTGAGGAATCTATAGGGCTTCAATTGTGGAAGGAAATAGCATTTGAAAAACAA 2433
QY 664 TTGAAAAGAAAAGGAAACATGAGATATTGACAGAGATTTATATGATGCTTTGAA 723
Db 2434 GAAAAGCATTTAAAGGAAAGTAAAGAAAAGCTGAGAGAGATCTCAATTTAAAGAAAG 2493
QY 724 ATACAGCTATTAGAACTTCATCAGAAATGAAAGTGTATTTATATACCAATCAATCT 783
Db 2494 GACATTGAGAAAGTGAAGAAAAGAAACCTTTGAAAAGAAAAGAAAGATTAACATGACAT 2553
QY 784 TCTTTACCTCAGACAAACAGAGGAATGTAGAGATTTCCAGGAATATCTTATATAGAA 843
Db 2554 AAGTCAGAAAAAGCAAAATTAAGATTTAGTGAATGTGTTATTAAGAAAAGAAAGGAC 2613
QY 844 AAAACAAATAGAGATTTATTAACAACAATGTTGAAGGAGAGAGGATATACATTAAGAA 903
Db 2614 AAGCTATATTCGCTACACAGAAAATGCCATTAAGAGAGTGAAGAGCAAAAATATCT 2673
QY 904 CATCTTGAAAGAAAAGAGATGTTCAATTAACCAAGAACAAAGAGATTAATCTGCT 963
Db 2674 GCTGCTATTAAAAAATGACAGACAGAGAAAAGTGAAGAAAAGATGATTAAGAAACAT 2733
QY 964 GACATTAACAAATCATCATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCAATA 1023
Db 2734 GACAAAGAAAAGCCTGAAAAAGAGA-----GGCATCTAGCAGAGAACCAAGAA 2781
QY 1024 AGTAAGTATGAGATGAATTAAGTCTGAATATGAGATTCATTAATTAAGTGAAGAAAG 1083
```

```
Db      2782 AAGCACTTGATGAGAAAAAATAACATCAGATTAATAGTAAATACGTAATCAGAA 2841
Qy      1084 GATGATGAAGACTTGACCAATTTAAAGCTTATGCAATATGACAAATTTCCAAAGTGA 1143
Db      2842 AAGGCAAAAAATMAAAAAAGACAGGAGCTAGATMAAAAGAAAAATCTAGAAATAAA 2901
Qy      1144 GAAAACATGAAATTTATTAAGAACTAGAAAGATTTGATGAGAAAAATGAAATTTAGAT 1203
Db      2902 GAAAGTATTAATTAATTAATCTACTCCAAACCATACGAGAGAAAAATCAAGTATAGTA 2961
Qy      1204 GATTAGATGAAGAAATAGAAAAATCATCAGAAATTAATCTGAGAAAAATTAAAAAA 1263
Db      2962 GACGGTAATTAAGCAACAATGAAAAACCTTATCCCTTAAGAAAAACAAAGATGAA 3021
Qy      1264 GGAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATTAAGTTGTAT 1323
Db      3022 CTTTGAAAACTCCAGATGAAAAAGAAAAAGATMAAAAGATTAATGATAGATAC 3081
Qy      1324 GATGACATATTTAAAAATATAAAAA 1349
Db      3082 AAGAAACGACAAACATTAAGTTAA 3107
```

```
RESULT 15
US-10-793-626-693
; Sequence 693, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 693
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-693
```

```
Query Match      4.2%; Score 62.8; DB 6; Length 627;
Best Local Similarity 48.6%; Pred. No. 0.41;
Matches 203; Conservative 0; Mismatches 212; Indels 3; Gaps 1;
```

```
Qy      1061 ATTCAATTAATGATGAAGAAAGATGATGAAGACTTAAAGCAATTTAAGCTATGTGC 1120
Db      134  ATCAAGTAAGTAAAGAAATTAACGAATTAAGAAAGAAAAAGCAAAATTTATTTAAAAAG 193
Qy      1121 AATATGACAATTTCCAAAGATGAAGAAAAACATGGAATTTATTAAGAACTAGAAATTTGA 1180
Db      194  TAAATGATTAAGTCAAGACACGTAAGAAAAACAGCTGAAGTATAGTTGAAAAATGTA 253
Qy      1181 TAGAGAAAAATGAAATTTAGATTTAGATGAAGAAATAGAAAAATCATCAGAAAGAT 1240
Db      254  AACAAAGACAAAGAAATTTGAAAAAGAGAGAGCTAGATTAATCTGAAAAAGCAT 313
Qy      1241 TATCTGAAGAAAAATTAAGAAAAAGAAATATGAAAAACAAAGATTAATTTTA 1300
Db      314  TTAACACAGCCAAAGCATATCTTGAAACATGTAGAAAAACAAAGCAAGAAAAAGAAAGTTG 373
Qy      1301 AACCAATGATTAAGTTGTATGATGACATATTAATAAATTAATAAATGATTAAGCAG 1360
Db      374  AACCACTTATGATGCTAT--TAAAGAAAAATATTAATCAATGATGCTTATGCAAG 430
Qy      1361 TTAATTAAGAAAAAGAAAAATTCATTAATTCATTTGTTTCAATATTTGACGAGACATG 1420
```

```
Db      431 CTTACAAAAAGCACTTAATTAAGAAAAAGAACTGTTTTCTTAATTGATGAAGATAATG 490
Qy      1421 AAATTTACAGATGCGATGATGATTATCTGAAGATATTAATTAATTTATGAAACT 1478
Db      491 CAACCAATTCGAAATGACGGAATATGAAAGATCTTTCTTAAGCATATTAAGAAAT 548
```

```
Search completed: December 7, 2005, 06:43:40
Job time : 598 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 01:48:14 Search time 1329 Seconds
(without alignments)
9221.383 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 1 1482 US-09-837-344-41

Sequence: 1 CAGAGACACACACGATCT.....ATATTTTATGAACTATA 1482

Scoring table: IDENTITY NUC
Gapop 10'-0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA Main.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1482	100.0	1482	3	US-09-837-344-41
2	1482	100.0	1482	3	US-09-837-344-41
3	1480.4	99.9	1482	3	US-09-837-344-45
4	1480.4	99.9	1482	3	US-09-900-963-45
5	1446	97.6	1493	3	US-09-837-344-38
6	1446	97.6	1493	3	US-09-900-963-32
7	950	64.1	950	3	US-09-837-344-32
8	950	64.1	950	3	US-09-900-963-32
9	472.8	31.9	988	3	US-09-837-344-34
10	472.8	31.9	988	3	US-09-900-963-34
11	470.8	31.8	954	3	US-09-837-344-37
12	470.8	31.8	954	3	US-09-900-963-37
13	233.4	15.7	1300	7	US-10-668-749A-1
14	153.2	10.3	1297	7	US-10-425-115-67687
15	151.2	10.2	6292	7	US-10-221-714A-661
16	149.2	10.1	6767	7	US-10-221-714A-330
17	146	9.9	3662	10	US-11-097-143-5152
18	145.8	9.8	1243	8	US-10-425-115-172717
19	145	9.8	1204	8	US-10-437-963-77858
20	142.4	9.6	1081	8	US-10-425-115-16756
21	139.4	9.4	9539	6	US-10-239-676-52
22	139.4	9.4	9539	6	US-10-240-453-54
23	137.8	9.3	6668	6	US-10-311-455-1670

24	132.8	9.0	778	8	US-10-363-345A-2179	Sequence 2179, Ap
25	132.8	9.0	778	8	US-10-363-345A-2180	Sequence 2180, Ap
26	132.8	9.0	778	9	US-10-363-483A-2179	Sequence 2179, Ap
27	132.8	9.0	778	9	US-10-363-483A-2180	Sequence 2180, Ap
28	131.4	8.9	7442	7	US-10-221-714A-409	Sequence 409, App
29	129.8	8.8	921	8	US-10-425-115-38710	Sequence 38710, A
30	129.8	8.8	1762	8	US-10-363-345A-7529	Sequence 7529, Ap
31	129.8	8.8	1762	8	US-10-363-345A-7530	Sequence 7530, Ap
32	129.8	8.8	1762	9	US-10-363-483A-7529	Sequence 7529, Ap
33	129.8	8.8	1762	9	US-10-363-483A-7530	Sequence 7530, Ap
34	129.2	8.7	1995	8	US-10-363-345A-7533	Sequence 7533, Ap
35	129.2	8.7	1995	8	US-10-363-345A-7534	Sequence 7534, Ap
36	129.2	8.7	1995	9	US-10-363-483A-7533	Sequence 7533, Ap
37	129.2	8.7	1995	9	US-10-363-483A-7534	Sequence 7534, Ap
38	129	8.7	1200	8	US-10-363-345A-25801	Sequence 25801, A
39	129	8.7	1200	8	US-10-363-345A-25802	Sequence 25802, A
40	129	8.7	1200	9	US-10-363-483A-25801	Sequence 25801, A
41	129	8.7	1200	9	US-10-363-483A-25802	Sequence 25802, A
42	129	8.7	7597	6	US-10-311-455-986	Sequence 986, App
43	129	8.7	14006	6	US-10-311-455-1931	Sequence 1931, Ap
44	128.6	8.7	3673778	6	US-10-312-841-1	Sequence 1, Appl
45	128.6	8.7	3673778	6	US-10-312-841-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-837-344-41
Sequence 41, Application US/09837344
Patent No. US2002004182A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

```
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-837-344-41

Query Match      100.0%; Score 1482; DB 3; Length 1482;
Best Local Similarity 100.0%; Pred. No. 2,4e-196;
Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGAACCAAAAGCGATCTAGAACAAAGAGACGCTTAAAGAAAGTTGCAAGAACAA 60
   |||
Db 1 CAAGAACCAAAAGCGATCTAGAACAAAGAGACGCTTAAAGAAAGTTGCAAGAACAA 60
QY 61 CAAGCGATTTAGAACAAAGATAGACTTGTAAAGAAAGTTACAGAGCAGCAAGCGAT 120
   |||
Db 61 CAAGCGATTTAGAACAAAGATAGACTTGTAAAGAAAGTTACAGAGCAGCAAGCGAT 120
QY 121 TTAGAACAAAGAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
   |||
Db 121 TTAGAACAAAGAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
QY 181 GAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGT 240
   |||
Db 181 GAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGT 240
QY 241 GCTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGATAGACTTGTAAAGAA 300
   |||
Db 241 GCTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGATAGACTTGTGTAAAGAA 300
QY 301 AAGTTACAAAGAGCAGCAAGCGATTTAGAACAGAGAGCGTCTAAAGAAAGTTGCAA 360
   |||
Db 301 AAGTTACAAAGAGCAGCAAGCGATTTAGAACAGAGAGCGTCTAAAGAAAGTTGCAA 360
QY 361 GAAACCAAAAGCGATTTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAACAA 420
   |||
Db 361 GAAACCAAAAGCGATTTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAACAA 420
QY 421 AGCGATTTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAACCAAGCGATTTA 480
   |||
Db 421 AGCGATTTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAACCAAGCGATTTA 480
QY 481 GAAACAAAGAGAGCGTCTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGAG 540
   |||
Db 481 GAAACAAAGAGAGCGTCTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGAG 540
QY 541 AGACGCTTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGAGAGCGTGT 600
   |||
Db 541 AGACGCTTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGAGAGCGTGT 600
QY 601 AAAGAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGCGTATCGAAGAAAA 660
   |||
Db 601 AAAGAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGCGTATCGAAGAAAA 660
QY 661 AATTAGAAAGAAAGAAAGCAATGAGATATTTAGAGAGAGATTTATATGTGCTTTA 720
   |||
Db 661 AATTAGAAAGAAAGAAAGCAATGAGATATTTAGAGAGAGATTTATATGTGCTTTA 720
QY 721 GAAATCCGAGCTTAAAGCTTCATCAGAAAGAAAGAGATTTATATTAACACATCAA 780
   |||
Db 721 GAAATCCGAGCTTAAAGCTTCATCAGAAAGAAAGAGATTTATATTAACACATCAA 780
QY 781 TCTTCTTACCTCAGAGCAACAGAGGAAATAGAGAGATTTCCAGAGAAATATCTATAATA 840
   |||
Db 781 TCTTCTTACCTCAGAGCAACAGAGGAAATAGAGAGATTTCCAGAGAAATATCTATAATA 840
QY 841 GAAAAAACAATAGAGATCTATTAACAACAAATGTTGAAGAGCAAGGATATACATTA 900
   |||
Db 841 GAAAAAACAATAGAGATCTATTAACAACAAATGTTGAAGAGCAAGGATATACATTA 900
```

```
QY 901 GCACATCTTGAAAGAAAGAAAGATGTTCAATTAACCAAGAACAAAGAGATTAATCT 960
   |||
Db 901 GCACATCTTGAAAGAAAGAAAGATGTTCAATTAACCAAGAACAAAGAGATTAATCT 960
QY 961 GCTGACATACAAATCATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCAA 1020
   |||
Db 961 GCTGACATACAAATCATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCAA 1020
QY 1021 ATTAAGTATGATGAGATGAAGAAATAGCTGTAATATGACATTCATTAATAGATAGAGA 1080
   |||
Db 1021 ATTAAGTATGATGAGATGAAGAAATAGCTGTAATATGACATTCATTAATAGATAGAGA 1080
QY 1081 GAAGATGATGAGACCTTAGAGCAATTTAAGCTATTTGCAATATGACAAATTTCCAGAT 1140
   |||
Db 1081 GAAGATGATGAGACCTTAGAGCAATTTAAGCTATTTGCAATATGACAAATTTCCAGAT 1140
QY 1141 GAAGAAACATAGCAATTTATTAAGACCTAGAGATTTGATAGAGAAAAATGAAATTTA 1200
   |||
Db 1141 GAAGAAACATAGCAATTTATTAAGACCTAGAGATTTGATAGAGAAAAATGAAATTTA 1200
QY 1201 GATGATTTAGATGAGAGATGAGAAATCATCAGAGAAATTTCTGAGAGAAAAATTA 1260
   |||
Db 1201 GATGATTTAGATGAGAGATGAGAAATCATCAGAGAAATTTCTGAGAGAAAAATTA 1260
QY 1261 AAAGAAAGAAATATGAGAAACCAAGAGATTAATTTTAAACCAATGATTAAGATTG 1320
   |||
Db 1261 AAAGAAAGAAATATGAGAAACCAAGAGATTAATTTTAAACCAATGATTAAGATTG 1320
QY 1321 TATGATGACATATTAATAATAATATTAATAGCAGTTAATAGAGAAAGGAAAA 1380
   |||
Db 1321 TATGATGACATATTAATAATAATATTAATAGCAGTTAATAGAGAAAGGAAAA 1380
QY 1381 TTCATTAATATCATGTTTCATATATTTAGACGAGACATGAAATTTACAGATCGTGAT 1440
   |||
Db 1381 TTCATTAATATCATGTTTCATATATTTAGACGAGACATGAAATTTACAGATCGTGAT 1440
QY 1441 GAGTATCTGAAGATATTAATTAATTTTATGAACATATA 1482
   |||
Db 1441 GAGTATCTGAAGATATTAATTAATTTTATGAACATATA 1482

RESULT 2
US-09-900-963-41
; Sequence 41, Application US/09900963
; Publication No. US20030064075A1
;
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUIDHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; OF STIMULATING THE T LYMPHOCYTES
;
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,963
; FILING DATE: 10-Jul-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/098,327
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
```

REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-900-963-41

Query Match 100.0%; Score 1482; DB 3; Length 1482;
Best Local Similarity 100.0%; Pred. No. 2,4e-196;
Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAGACAAACGCGATCTAGAAACAAGAGACGCTCTAAAGAAAAGTTGCAAGAACAA 60
DB 1 CAAAGACAAACGCGATCTAGAAACAAGAGACGCTCTAAAGAAAAGTTGCAAGAACAA 60
QY 61 CAAAGCGATTTAGAACAAAGTAGAGCTGCTAAAGAAAAGTTGCAAGAACGCGAT 120
DB 61 CAAAGCGATTTAGAACAAAGTAGAGCTGCTAAAGAAAAGTTGCAAGAACGCGAT 120
QY 121 TTAGAACAGAGACCTTGTAAAGAAAAGTTGCAAGAACAAACGCGATCTAGAACAA 180
DB 121 TTAGAACAGAGACCTTGTAAAGAAAAGTTGCAAGAACAAACGCGATCTAGAACAA 180
QY 181 GAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAGAGACGCT 240
DB 181 GAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAGAGACGCT 240
QY 241 GCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAGAGACGCTGCTAAAGAAA 300
DB 241 GCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAGAGACGCTGCTAAAGAAA 300
QY 301 AAGTTACAGAGACGCGATTTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAAG 360
DB 301 AAGTTACAGAGACGCGATTTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAAG 360
QY 361 GAACAAACAAAGCGATTTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAA 420
DB 361 GAACAAACAAAGCGATTTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAA 420
QY 421 AGCGATTTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTA 480
DB 421 AGCGATTTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTA 480
QY 481 GAACAAAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAGAG 540
DB 481 GAACAAAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAGAG 540
QY 541 AGACGCTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAGAGACGCTGCT 600
DB 541 AGACGCTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAGAGACGCTGCT 600
QY 601 AAAAGAAAAGTTGCAAGAGCAGCAAGATTTAGAACAAAGAGCGCTGATACGAAAAA 660
DB 601 AAAAGAAAAGTTGCAAGAGCAGCAAGATTTAGAACAAAGAGCGCTGATACGAAAAA 660
QY 661 AATTAGAAAAGAAAAGCAATGAGATATATTAGACAGAGATTTATATGTCGCTTTA 720
DB 661 AATTAGAAAAGAAAAGCAATGAGATATATTAGACAGAGATTTATATGTCGCTTTA 720

QY 721 GAATATACAGCTATAGAACTTCATCAGAAAAATGAACGTGATATTATATACCATCA 780
DB 721 GAATATACAGCTATAGAACTTCATCAGAAAAATGAACGTGATATTATATACCATCA 780
QY 781 TCTTCTTTTACCTCAGAGCAACAGAGGAAATAGTAGATTCCAGAAATATCTATATATA 840
DB 781 TCTTCTTTTACCTCAGAGCAACAGAGGAAATAGTAGATTCCAGAAATATCTATATATA 840
QY 841 GAAAAAACAATAGAAATCTTATTAACAACAAAGTTGAAGAGAGAGGATATACATPAA 900
DB 841 GAAAAAACAATAGAAATCTTATTAACAACAAAGTTGAAGAGAGAGATATACATPAA 900
QY 901 GGACATCTTGAAGAAAAAGAGATGTTCAATATAACAGAACAAAGAGATTAATCT 960
DB 901 GGACATCTTGAAGAAAAAGAGATGTTCAATATAACAGAACAAAGAGATTAATCT 960
QY 961 GCTGACATACAAAATCATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCAA 1020
DB 961 GCTGACATACAAAATCATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCAA 1020
QY 1021 ATAAAGTAAAGTATAGAGATGAATTAAGTCTGAATATGACATTCATTAATAGTGAAGA 1080
DB 1021 ATAAAGTAAAGTATAGAGATGAATTAAGTCTGAATATGACATTCATTAATAGTGAAGA 1080
QY 1081 GAAGATGATGAGACCTTAGACGAATTTAAGCCTATTGTGCAATATGACAAATTTCCAAGAT 1140
DB 1081 GAAGATGATGAGACCTTAGACGAATTTAAGCCTATTGTGCAATATGACAAATTTCCAAGAT 1140
QY 1141 GAAGAAAACATAGAAATTTATTAAGAACTGAAGATTTGATAGAGAAAAATGAATTTTA 1200
DB 1141 GAAGAAAACATAGAAATTTATTAAGAACTGAAGATTTGATAGAGAAAAATGAATTTTA 1200
QY 1201 GATGATTTAGATGAGGAATAGAAAAATCATCGAAGAAATTAATCTGAAAGAAAAATPAAA 1260
DB 1201 GATGATTTAGATGAGGAATAGAAAAATCATCGAAGAAATTAATCTGAAAGAAAAATPAAA 1260
QY 1261 AAAGAAAGAAATATGAAAAAACAAGAGATTAATTTTAAACCAATGATTAAGATTTG 1320
DB 1261 AAAGAAAGAAATATGAAAAAACAAGAGATTAATTTTAAACCAATGATTAAGATTTG 1320
QY 1321 TATGATGAGCATATTAATAAATATTAATAATGATTAAGAGGTTATTAAGAAAAAGAAAA 1380
DB 1321 TATGATGAGCATATTAATAAATATTAATAATGATTAAGAGGTTATTAAGAAAAAGAAAA 1380
QY 1381 TTGATTAATCATTTGTTCTATATTTTGAACGAGACAAAGAAATTTTACAGATCGTGAT 1440
DB 1381 TTGATTAATCATTTGTTCTATATTTTGAACGAGACAAAGAAATTTTACAGATCGTGAT 1440
QY 1441 GAGTTATCTGAAGATATTAATAATATTTATGAACTATPAA 1482
DB 1441 GAGTTATCTGAAGATATTAATAATATTTATGAACTATPAA 1482

RESULT 3

US-09-837-344-45
Sequence 45, Application US/09837344
Patent No. US2002004182A1

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine

DRUIHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE

HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE

OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P. O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/837,344
APPLICATION NUMBER: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-837-344-45

Query Match 99.9%; Score 1480.4; DB 3; Length 1482;
Best Local Similarity 99.9%; Pred. No. 4e-196;
Matches 1481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CAAGAACAAAGCATCTAGAACAAAGACGCTTAAGAAAGTTGCAAGAACAA 60
1 CAAGAACAAAGCATCTAGAACAAAGACGCTTAAGAAAGTTGCAAGAACAA 60
61 CAAGGATTTAGAACAAAGATCTAGAACAAAGACGCTTAAGAAAGTTGCAAGAACAA 120
61 CAAGGATTTAGAACAAAGATCTAGAACAAAGACGCTTAAGAAAGTTGCAAGAACAA 120
121 TTAGAACAAAGACGCTTAAGAACAAAGATCTAGAACAAAGACGCTTAAGAACAA 180
121 TTAGAACAAAGACGCTTAAGAACAAAGATCTAGAACAAAGACGCTTAAGAACAA 180
181 GAGAGCGCTTAAGAACAAAGATCTAGAACAAAGACGCTTAAGAACAAAGACGCT 240
181 GAGAGCGCTTAAGAACAAAGATCTAGAACAAAGACGCTTAAGAACAAAGACGCT 240
181 GAGAGCGCTTAAGAACAAAGATCTAGAACAAAGACGCTTAAGAACAAAGACGCT 240
241 GCTTAAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGAT 300
241 GCTTAAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGAT 300
301 AAGTTAAGACGACGCAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAA 360
301 AAGTTAAGACGACGCAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAA 360
361 GAACAAAGACGATTTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCT 420
361 GAACAAAGACGATTTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCT 420
421 AGGATTTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGAT 480
421 AGGATTTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGAT 480
481 GAACAAAGACGATTTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCT 540
481 GAACAAAGACGATTTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCT 540

481 GAACAAAGACGATTTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCT 540
541 AGGATTTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGAT 600
541 AGGATTTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGAT 600
601 AAAGAAAGTTGCAAGACGCAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAA 660
601 AAAGAAAGTTGCAAGACGCAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAA 660
661 AATTAGAAAGAAAGAAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGAT 720
661 AATTAGAAAGAAAGAAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGAT 720
721 GAATACCAAGCTATAGAACTTCCATCAGAAATAGAACGATATATATATATATATATAT 780
721 GAATACCAAGCTATAGAACTTCCATCAGAAATAGAACGATATATATATATATATATAT 780
781 TCTTCTTACCTCAGACCAAGACGATATAGAACGATATATATATATATATATATATAT 840
781 TCTTCTTACCTCAGACCAAGACGATATAGAACGATATATATATATATATATATATAT 840
841 GAAAAAACAATAGAAATCTATTAACAACAATGTTGAAGACGAAAGGATATATATATAT 900
841 GAAAAAACAATAGAAATCTATTAACAACAATGTTGAAGACGAAAGGATATATATATAT 900
901 GGACATCTTGAAGAAAGAAAGATGTTCAATTAACAACAATGTTGAAGATATATATAT 960
901 GGACATCTTGAAGAAAGAAAGATGTTCAATTAACAACAATGTTGAAGATATATATATAT 960
961 GCTGACATACAAATCATATACATTAAGACAGTAATATATATATATATATATATATAT 1020
961 GCTGACATACAAATCATATACATTAAGACAGTAATATATATATATATATATATATAT 1020
1021 AATAAGTATAGAGATGAAGAAATAGCTGATATATATATATATATATATATATATAT 1080
1021 AATAAGTATAGAGATGAAGAAATAGCTGATATATATATATATATATATATATATAT 1080
1081 GAAGATGATGAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCT 1140
1081 GAAGATGATGAAGATCTAGAACAAAGATCTAGAACAAAGATCTAGAACAAAGATCT 1140
1141 GAAGAAACATAGAAATTTATTAAGAACTAGAACAAAGATCTAGAACAAAGATCT 1200
1141 GAAGAAACATAGAAATTTATTAAGAACTAGAACAAAGATCTAGAACAAAGATCT 1200
1201 GATGATTTAGTGAAGAAATGAAATATATATATATATATATATATATATATATAT 1260
1201 GATGATTTAGTGAAGAAATGAAATATATATATATATATATATATATATATATATAT 1260
1261 AAAGAAAGAAATATGAAGAAACAAAGATATATATATATATATATATATATATATAT 1320
1261 AAAGAAAGAAATATGAAGAAACAAAGATATATATATATATATATATATATATATAT 1320
1321 TATGATGACATTTATTAAGAAATATATATATATATATATATATATATATATATAT 1380
1321 TATGATGACATTTATTAAGAAATATATATATATATATATATATATATATATATATAT 1380
1381 TTGATTAAT 1440
1381 TTGATTAAT 1440
1441 GAGTATGCAAGAT 1482
1441 GAGTATGCAAGAT 1482

RESULT 4
US-09-900-963-45
Sequence 45, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine

DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/900,963
APPLICATION NUMBER: 08/098,327
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-900-963-45
Query Match 99.9%; Score 1480.4; DB 3; Length 1482;
Best Local Similarity 99.9%; Pred. No. 4e-196;
Matches 1481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	CAAGAACAAACGATCTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA	60
Db	1	CAAGAACAAACGATCTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA	60
Qy	61	CAAGCGATTAGAACAGATGAGCTTGCTAAAGAAAAGTTACAAAGACCAAAAGCGAT	120
Db	61	CAAGCGATTAGAACAGATGAGCTTGCTAAAGAAAAGTTACAAAGACCAAAAGCGAT	120
Qy	121	TTAAGAACAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATCTGAACAA	180
Db	121	TTAAGAACAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATCTGAACAA	180
Qy	181	GAGAGACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCT	240
Db	181	GAGAGACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCT	240
Qy	241	GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATGAGCTTGCTAAAGAA	300
Db	241	GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATGAGCTTGCTAAAGAA	300
Qy	301	AAGTTACAGAGCGCAAAAGGATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCA	360

Db	301	AAGTTACAGAGCGCAAAAGGATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCA	360
Qy	361	GAACAAACAAAGCATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA	420
Db	361	GAACAAACAAAGCATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA	420
Qy	421	AGCGATTTAGAACAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA	480
Db	421	AGCGATTTAGAACAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA	480
Qy	481	GAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG	540
Db	481	GAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG	540
Qy	541	AGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGTCT	600
Db	541	AGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGTCT	600
Qy	601	AAAGAAAAGTTGCAAGAGAGCGCAAAAGGATTTAGAACAAAGAGCGTATACGAAAAA	660
Db	601	AAAGAAAAGTTGCAAGAGAGCGCAAAAGGATTTAGAACAAAGAGCGTATACGAAAAA	660
Qy	661	AATTTAGAAAAGAAAAGAGAACATGAGATATATAGCAGAGATTTATATGTCGTTTA	720
Db	661	AATTTAGAAAAGAAAAGAGAACATGAGATATATAGCAGAGATTTATATGTCGTTTA	720
Qy	721	GAATATACAGCTATAGAACTTCCATCAGAAAATGAAACGTGATATTTATACCATCAA	780
Db	721	GAATATACAGCTATAGAACTTCCATCAGAAAATGAAACGTGATATTTATACCATCAA	780
Qy	781	TCTTCTTTTACCTCAGGCAACAGAGGAAATGATGAGATTCCAAAGGAAATATCTATATA	840
Db	781	TCTTCTTTTACCTCAGGCAACAGAGGAAATGATGAGATTCCAAAGGAAATATCTATATA	840
Qy	841	GAAGAAAACAAATAGAGATCTATTACAAACAAATGTTGAAGAGAGGATATACATATA	900
Db	841	GAAGAAAACAAATAGAGATCTATTACAAACAAATGTTGAAGAGAGGATATACATATA	900
Qy	901	GGAATCTTGAAGAAAAGAAAGATGTTCAATATAAACAGAACAAAGAAAGTAAATCT	960
Db	901	GGAATCTTGAAGAAAAGAAAGATGTTCAATATAAACAGAACAAAGAAAGTAAATCT	960
Qy	961	GCTGACATCAAAATCATCATTTAGAGACAGTAAATTTCTGATGTTAATGATTTCAA	1020
Db	961	GCTGACATCAAAATCATCATTTAGAGACAGTAAATTTCTGATGTTAATGATTTCAA	1020
Qy	1021	ATAAGTAAATGATGAGATGAAATGAGTCTGAATATGACGATTCATTAATGATGAAGA	1080
Db	1021	ATAAGTAAATGATGAGATGAAATGAGTCTGAATATGACGATTCATTAATGATGAAGA	1080
Qy	1081	GAAGATGATGAAGCTTTAGACGAATTTTAAAGCTTATGTCGAATTTCCAGAT	1140
Db	1081	GAAGATGATGAAGCTTTAGACGAATTTTAAAGCTTATGTCGAATTTCCAGAT	1140
Qy	1141	GAAGAAAACATAGAAATTTTAAAGAACTGGAAGATTTGATAGAGAAAATGAAAATTTA	1200
Db	1141	GAAGAAAACATAGAAATTTTAAAGAACTGGAAGATTTGATAGAGAAAATGAAAATTTA	1200
Qy	1201	GATGATTTAGATGAAGAAATGAGAAATTCATCGAAGAAATTTATCTGAGAAAATATAA	1260
Db	1201	GATGATTTAGATGAAGAAATGAGAAATTCATCGAAGAAATTTATCTGAGAAAATATAA	1260
Qy	1261	AAAGAAAAGAAATGAGAAAACAAAGATTAATTTTAAACCAATGATTAAGATTTTG	1320
Db	1261	AAAGAAAAGAAATGAGAAAACAAAGATTAATTTTAAACCAATGATTAAGATTTTG	1320
Qy	1321	TATGATGAGATTTTAAATAATTAATAATGATAGAGGTTATATAGAAAAGGAAAAA	1380
Db	1321	TATGATGAGATTTTAAATAATTAATAATGATAGAGGTTATATAGAAAAGGAAAAA	1380
Qy	1381	TTGATTAATCATTTGTTTCAATATTTTGAAGAGAGCAATGAAATTTTACGATGTCGAT	1440
Db	1381	TTGATTAATCATTTGTTTCAATATTTTGAAGAGAGCAATGAAATTTTACGATGTCGAT	1440

QY 1441 GAGTATCTGAAGATATTAATAATTTTATGAACCTATA 1482
DB 1441 GAGTATCTGAAGATATTAATAATTTTATGAACCTATA 1482

RESULT 5
US-09-837-344-38
Sequence 38, Application US/09837344
Patent No. US2002004182A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRULHE, Pierrie
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STRGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-837-344-38

Query Match 97.6%; Score 1446; DB 3; Length 1493;
Best Local Similarity 99.8%; Pred. No. 2.4e-191;
Matches 1479; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 1 CAAGAACCAAAAGCATCTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACAA 60
DB 1 CAAGAACCAAAAGCATCTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACAA 60

QY 61 CAAGGATTTTGAACAAGTAGACTTGCTAAAGAAAAGTTACAAGAGCAAAAGCGAT 120
DB 61 CAAGGATTTTGAACAAGTAGACTTGCTAAAGAAAAGTTACAAGAGCAAAAGCGAT 120

QY 121 TTAGAACAAGAGACTTGCTAAAGAAAAGTTGCAAGAACCAAAAGCGATCTAGAACAA 180
DB 121 TTAGAACAAGAGACTTGCT-AAAAGAAAAGTTGCAAGAACCAAAAGCGATCTAGAACAA 179

QY 181 GAGAGACGTGCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAGAGAGCGT 240
DB 180 GAGAGACGTGCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAGAGAGCGT 239

QY 241 GCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAGATAGA CTGGCTAAAGAA 300
DB 240 GCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAGATAGA CTGGCTAAAGAA 299

QY 301 AAGTTACAAAGACGACCAAAAGCGATTTAGAACAGAGAACGCTCTAAAGAAAAGTTGCAA 360
DB 300 AAGTTACAAAGACGACCAAAAGCGATTTAGAACAGAGAACGCTCTAAAGAAAAGTTGCAA 359

QY 361 GAACAAACAAGCGATTTAGAACAGAGAGCGTGTAAAGAAAAGTTGCAAGAACCAACA 420
DB 360 GAACAAACAAGCGATTTAGAACAGAGAGCGTGT-AAAAGAAAAGTTGCAAGAACCAACA 418

QY 421 AGCGATTTAGAACAGAGAGACTTGCTTAAGAAAAGTTGCAAGAACCAAAAGCGATTTA 480
DB 419 AGCGATTTAGAACAGAGAGACTTGCTTAAGAAAAGTTGCAAGAACCAAAAGCGATTTA 478

QY 481 GAACAGAGAGACGCTGCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAGAG 540
DB 479 GAACAGAGAGACGCTGCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAGAG 538

QY 541 AGAGCTGCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAGAGAGAGCTGCT 600
DB 539 AGAGCTGCT-AAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAGAGAGAGCTGCT 597

QY 601 AAAAGAAAAGTTGCAAGAGACGCAAAAGAGATTTAGAACAAAGAGAGCTGATACGAAAAA 660
DB 598 AAAAGAAAAGTTGCAAGAGACGCAAAAGAGATTTAGAACAAAGAGAGCTGATACGAAAAA 657

QY 661 AATTAGAAAAGAAAAGAAAACATGAGATATTTAGCAGAGATTTATATGCTGTTA 720
DB 658 AATTAGAAAAGAAAAGAAAACATGAGATATTTAGCAGAGATTTATATGCTGTTA 717

QY 721 GAATATACGAGCATATGAACTTCCATCAGAAAATGAAAGCTGATATTTATATACCATCA 780
DB 718 GAATATACGAGCATATGAACTTCCATCAGAAAATGAAAGCTGATATTTATATACCATCA 777

QY 781 TCTTCTTTTACCTCAGAGAACAGAGGAAATGTAAGATTCCAGAAAATATCTATATA 840
DB 778 TCTTCTTTTACCTCAGAGAACAGAGGAAATGTAAGATTCCAGAAAATATCTATATA 837

QY 841 GAAAAAACAATATGAGATCTATTACAACAATGTTGAAGACGAAAGGATATTAATTA 900
DB 838 GAAAAAACAATATGAGATCTATTACAACAATGTTGAAGACGAAAGGATATTAATTA 897

QY 901 GGACATCTTGAAGAAAAGAAAGATGCTTCAATPAAACAGAACCAAAAAGAAATATCT 960
DB 898 GGACATCTTGAAGAAAAGAAAGATGCTTCAATPAAACAGAACCAAAAAGAAATATCT 957

QY 961 GCTGACATACAAATCATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTCAA 1020
DB 958 GCTGACATACAAATCATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTCAA 1017

QY 1021 AATAAGTATGATGAGATGAATAATAAGTGTGAATATGACGATTCATTAATAGTAAGAA 1080
DB 1018 AATAAGTATGATGAGATGAATAATAAGTGTGAATATGACGATTCATTAATAGTAAGAA 1077

QY 1081 GAAGATGATGAAGACTTTGAGCAATTTAAAGCTATTTGCAATATGACAAATTTCCAGAT 1140
DB 1078 GAAGATGATGAAGACTTTGAGCAATTTAAAGCTATTTGCAATATGACAAATTTCCAGAT 1137

QY 1141 GAAGAAAACATAGAAATTTATPAAAGAACTAGAAAGATTGATAGAGAAAAATGAAATTTA 1200
DB 1138 GAAGAAAACATAGAAATTTATPAAAGAACTAGAAAGATTGATAGAGAAAAATGAAATTTA 1197

QY 1201 GATGATTTTATGATGAAGAAATAGAAAATCATAGAAAGATTTATCTGAAGAAAAATPAAA 1260
DB 1198 GATGATTTTATGATGAAGAAATAGAAAATCATAGAAAGATTTATCTGAAGAAAAATPAAA 1257

1141 GAAGAAACATGGAATTTATTAAGAACTGAGAGATTTGATGAGAAAAATGAATAATTTA 1200
1138 GAAGAAACATGGAATTTATTAAGAACTGAGAGATTTGATGAGAAAAATGAATAATTTA 1197
1201 GATGATTTGATGAGAAATGAAAAATCATCAGAAATTTCTGAGAAAAATTAATA 1260
1198 GATGATTTGATGAGAAATGAAAAATCATCAGAAATTTCTGAGAAAAATTAATA 1257
1261 AAAGGAAAGAAATGAAAAACAAAGATTAATTTTAAACCAATGATTAAGTTTG 1320
1258 AAAGGAAAGAAATGAAAAACAAAGATTAATTTTAAACCAATGATTAAGTTTG 1317
1321 TATGATGAGCAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
1318 TATGATGAGCAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1377
1381 TTCATTAATCATTTGTTTATATTTTGAAGCAATGAAATTTTACAGATCGTGAT 1440
1378 TTCATTAATCATTTGTTTATATTTTGAAGCAATGAAATTTTACAGATCGTGAT 1437
1441 GAGTATCTGAAGATTAATACTAAATTTTATGAAACTATTA 1482
1438 GAGTATCTGAAGATTAATACTAAATTTTATGAAACTATTA 1479

RESULT 7
US-09-837-344-32
Sequence 32, Application US/09837344
Patent No. US2002004182A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-837-344-32
Query Match 64.1%; Score 950; DB 3; Length 950;
Best Local Similarity 100.0%; Pred. No. 1,2e-122;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

11 AAAGGATCTGAAACAAAGAGACGCTGCTAAAGAAAACCTTCCAGAAACAAACCGGATT 70
1 AAAGGATCTGAAACAAAGAGACGCTGCTAAAGAAAACCTTCCAGAAACAAACCGGATT 60
71 TAGAACAAGATAGACTTCTTAAAGAAAAGTTACAAAGACAGCAAGCGATTAGAACAG 130
61 TAGAACAAGATAGACTTCTTAAAGAAAAGTTACAAAGACAGCAAGCGATTAGAACAG 120
121 AGAGACTTGCTTAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAAAGAGACGCTG 180
131 AGAGACTTGCTTAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAAAGAGACGCTG 190
121 AGAGACTTGCTTAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAAAGAGACGCTG 180
191 CTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGACGCTGTAAGAAA 250
181 CTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGACGCTGTAAGAAA 240
251 AGTTCAGAGAACAAAGAGCGATTTTGAACAAAGATAGACTTGTCTAAAGAAAAGTTACAAG 310
241 AGTTCAGAGAACAAAGAGCGATTTTGAACAAAGATAGACTTGTCTAAAGAAAAGTTACAAG 300
311 AGCAGCAAGAGCGATTTTGAACAAAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAA 370
301 AGCAGCAAGAGCGATTTTGAACAAAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAA 360
371 GCGATTTGAAACAAAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGCTTGA 430
361 GCGATTTGAAACAAAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGCTTGA 420
431 AACAGAGAGACTTCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAG 490
421 AACAGAGAGACTTCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAG 480
491 GACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGACGCTGTA 550
481 GACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGACGCTGTA 540
551 AAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGACGCTGTAAGAAAAGT 610
541 AAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGACGCTGTAAGAAAAGT 600
611 TGCAAGAGCAGCAAGAGATTTTGAACAAAGAGCGTGAACGAAAAAATTTAGAAA 670
601 TGCAAGAGCAGCAAGAGATTTTGAACAAAGAGCGTGAACGAAAAAATTTAGAAA 660
671 GAAAAAAGGAAACATGAGATATATTAGCAGAGATTTATATGCTGTTAGAAATACAG 730
661 GAAAAAAGGAAACATGAGATATATTAGCAGAGATTTATATGCTGTTAGAAATACAG 720
731 CTATGAACCTTCATCAGAAAATGAAACGATTAATTAATTAATTAATTAATTAATTA 790
721 CTATGAACCTTCATCAGAAAATGAAACGATTAATTAATTAATTAATTAATTAATTA 780
791 CTCAGAGCAACAGAGGAAATAGTAGAATTCAGAGAAATATCTAATATAGAAAAACA 850
781 CTCAGAGCAACAGAGGAAATAGTAGAATTCAGAGAAATATCTAATATAGAAAAACA 840
851 ATAGAGAAATCTATTAACAACAAATGTTGAAGAGCAAGAGGATATACATAAAGACATCTTG 910
841 ATAGAGAAATCTATTAACAACAAATGTTGAAGAGCAAGAGGATATACATAAAGACATCTTG 900
911 AAAGAAAAGAAATGTTCAATTAACCAAGAACAAAGAAAGAAATTAATCT 960
901 AAAGAAAAGAAATGTTCAATTAACCAAGAACAAAGAAAGAAATTAATCT 950

RESULT 8
US-09-900-963-32
Sequence 32, Application US/0900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-900-963-32
Query March 64.1%; Score 950; DB 3; Length 950;
Best Local Similarity 100.0%; Pred. No. 1,2e-122; Mismatches 0; Indels 0; Gaps 0;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 AAAGCGATCTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAACAAAGCGATT 70
DB 1 AAAGCGATCTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAACAAAGCGATT 60
QY 71 TAGAACAGATAGACTTCTTAAAGAAAAGTTTCAAGAGACGCAAGCGATTTAACAAAG 130
DB 61 TAGAACAGATAGACTTCTTAAAGAAAAGTTTCAAGAGACGCAAGCGATTTAACAAAG 120
QY 131 AGAGACTGTCTAAGAAAAGTTGCAAGAACAAACAAAGCGATCTAGAACAAAGAGAGACGTG 190
DB 121 AGAGACTGTCTAAGAAAAGTTGCAAGAACAAACAAAGCGATCTAGAACAAAGAGACGTG 180
QY 191 CTAAGAAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGACGTGCTTAAGAA 250
DB 181 CTAAGAAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGACGTGCTTAAGAA 240
QY 251 AGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGACGTGCTTAAGAAAAGTTACAG 310

DB 241 AGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGATAGACTGTCTAAGAAAAGTTACAG 300
QY 311 AGCAGCAAGCGATTTAGAACAAAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAAACAA 370
DB 301 AGCAGCAAGCGATTTAGAACAAAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAAACAA 360
QY 371 GCGATTTAGAACAAAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAAACAAAGCGATTAG 430
DB 361 GCGATTTAGAACAAAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAAACAAAGCGATTAG 420
QY 431 AACAGAGACACTGTCTAAGAAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGA 490
DB 421 AACAGAGACACTGTCTAAGAAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGA 480
QY 491 GACGTGCTAAGAAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGACGTGCTA 550
DB 481 GACGTGCTAAGAAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGACGTGCTA 540
QY 551 AAGAAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGACGTGCTTAAAGAAAAGT 610
DB 541 AAGAAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGACGTGCTTAAAGAAAAGT 600
QY 611 TGCAGAGCAGCAAGAGATTTAGAACAAAGAAAGCTGATACGAAAAAAATTTAGAAA 670
DB 601 TGCAGAGCAGCAAGAGATTTAGAACAAAGAAAGCTGATACGAAAAAAATTTAGAAA 660
QY 671 GAAAAAAGCAACTGGAGATATTTTGACAGAGATTTATATGTCGTTTGAATACAG 730
DB 661 GAAAAAAGCAACTGGAGATATTTTGACAGAGATTTATATGTCGTTTGAATACAG 720
QY 731 CTATAGAACTTCATCAGAAAAATGAAAGCTGATATATATACACATCTCTCTTAC 790
DB 721 CTATAGAACTTCATCAGAAAAATGAAAGCTGATATATATACACATCTCTCTTAC 780
QY 791 CTCAGGCAACAGAGGAAATAGTAGATTCAGAGAAATATATATATAGAAAAACAA 850
DB 781 CTCAGGCAACAGAGGAAATAGTAGATTCAGAGAAATATATATATAGAAAAACAA 840
QY 851 ATAGAGATTTATTAACAACAATGTTGAGAGACGAAAGGATATATACATTAAGACATCTTG 910
DB 841 ATAGAGATTTATTAACAACAATGTTGAGAGACGAAAGGATATATACATTAAGACATCTTG 900
QY 911 AAGAAAAGAAAGATGGTTCAATTAACCCAGAACAAAAGAGTAATCT 960
DB 901 AAGAAAAGAAAGATGGTTCAATTAACCCAGAACAAAAGAGTAATCT 950
RESULT 9
US-09-837-344-34
Sequence 34, Application US/09837344
Patent No. US2002004182A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001

```
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/462,625
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: FR 91 01286
/ FILING DATE: 05-FEB-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McGowan, Malcolm K.
/ REGISTRATION NUMBER: 39,300
/ REFERENCE/DOCKET NUMBER: 010830-078
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-2021
/ INFORMATION FOR SEQ ID NO: 34:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 988 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ PUBLICATION INFORMATION:
/ DOCUMENT NUMBER: WO 92/13884
/ PUBLICATION DATE: 20-AUG-1992
/ SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-837-344-34

Query Match      31.9%; Score 472.8; DB 3; Length 988;
Best Local Similarity 95.7%; Pred. No. 1.7e-56;
Matches 486; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 77 AAGATGACTTGGCTAAAGAAAGTTCAAGAGCAGCAAGCGATTAGAACAGAGAGAC 136
DB 481 AAGAAAACAGCTAAAGAAAGTTACAGGGCGCAAAAGCGATTCAAGAACAGAGAGAC 540
QY 137 TTGCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAG 196
DB 541 GTGCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAG 600
QY 197 AAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 256
DB 601 AAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 257 AAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 316
DB 661 AAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 317 AAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 376
DB 721 AAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 377 TGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
DB 781 TGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 437 AGAGACTTCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAG 496
DB 841 AGAGACTTCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAG 900
QY 497 CTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAG 556
DB 901 CTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAG 960
QY 557 AGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 584
DB 961 GGTTCAGAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 988
```

```
/ DRUILHE, Pierre
/ TITLE OF INVENTION: HEPATIC SEQUENCES SPECIFIC FOR THE
/ HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
/ OF STIMULATING THE T LYMPHOCYTES
/ NUMBER OF SEQUENCES: 46
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Burns, Doane, Swecker & Mathis
/ STREET: P.O. Box 1404
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: United States
/ ZIP: 22313-1404
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/900,963
/ FILING DATE: 10-Jul-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/098,327
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McGowan, Malcolm K.
/ REGISTRATION NUMBER: 39,300
/ REFERENCE/DOCKET NUMBER: 010830-045
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-2021
/ INFORMATION FOR SEQ ID NO: 34:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 988 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ PUBLICATION INFORMATION:
/ DOCUMENT NUMBER: WO 92/13884
/ PUBLICATION DATE: 20-AUG-1992
/ SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-900-963-34

Query Match      31.9%; Score 472.8; DB 3; Length 988;
Best Local Similarity 95.7%; Pred. No. 1.7e-56;
Matches 486; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 77 AAGATGACTTGGCTAAAGAAAGTTCAAGAGCAGCAAGCGATTAGAACAGAGAGAC 136
DB 481 AAGAAAACAGCTAAAGAAAGTTACAGGGCGCAAAAGCGATTCAAGAACAGAGAGAC 540
QY 137 TTGCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAG 196
DB 541 GTGCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAG 600
QY 197 AAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 256
DB 601 AAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 257 AAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 316
DB 661 AAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 317 AAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 376
DB 721 AAGCGATTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 377 TGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
DB 781 TGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 437 AGAGACTTCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAGAGAGAG 496
```

|||||
Db 841 AGAAGCTTCTAAAGAAAAGTTACAGAGCGCAAGCGATTTCAGAACAGATGACTTG 900
Qy 497 CTAAGAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTAAAGAA 556
Db 901 CTAAGAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTAAAGAA 960
Qy 557 AGTTGCAAGAACAAAGCGATTTAGA 584
Db 961 GGTGCAAGAACAAAGCGATTTAGA 988
RESULT 11
US-09-837-344-37
Sequence 37, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATITIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837.344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 954 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..954
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-09-837-344-37
Query Match 31.8%; Score 470.8; DB 3; Length 954;
Best Local Similarity 95.7%; Pred. No. 3.2e-56;
Matches 484; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 77 AAGTAGCTTCTAAAGAAAAGTTACAGAGCGCAAGCGATTTAGAACAGAGAGAC 136
|||||

Db 449 AAGAAAAGAGCGTAAAGAAAAGTTACAGGGGCAACAAGCGATTCAGAACAGAGAGAC 508
Qy 137 TTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTCAGAACAGAGAGCGTGCTAAAG 196
Db 509 GTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTCAGAACAGAGAGCGTGCTAAAG 568
Qy 197 AAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTAAAGAAAAGTTG 256
Db 569 AAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTAAAGAAAAGTTG 628
Qy 257 AAGAACAAAGCGATTTAGAACAGATGACTTGCTAAAGAAAAGTTACAGAGCAGC 316
Db 629 AAGAACAAAGCGATTTAGAACAGAGAGCGTGCTAAAGAAAAGTTGCAAGAAC 688
Qy 317 AAAGCGATTTAGAACAGAGAGCGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATT 376
Db 689 AAAGCGATTTAGAACAGAGAGCGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATT 748
Qy 377 TAGAACAGAGAGCGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAG 436
Db 749 TAGAACAGAGAGCGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAG 808
Qy 437 AGAAGCTTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTG 496
Db 809 AGAAGCTTCTAAAGAAAAGTTGCAAGAGCAGCAAGCGATTTAGAACAGATGACTTG 868
Qy 497 CTAAGAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTAAAGAA 556
Db 869 CTAAGAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTAAAGAA 928
Qy 557 AGTTGCAAGAACAAAGCGATTTA 582
Db 929 GGTGCAAGAACAAAGCGATTTA 954
RESULT 12
US-09-900-963-37
Sequence 37, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATITIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900.963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 37:

```
SEQUENCE CHARACTERISTICS:
LENGTH: 954 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..954
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-09-900-963-37
```

```
Query Match 31.8%; Score 470.8; DB 3; Length 954;
Best Local Similarity 95.7%; Pred. No. 3.2e-56;
Matches 484; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
```

```
QY 77 AAGATGACTGCTTAAGAAAGTTACAGAGACGCAAGCGATTTAGAACAGAGAGAC 136
DB 449 AAGAAAAAGCAGCTTAAGAAAGTTACAGAGGCAAGCGATTTAGAACAGAGAGAC 508
QY 137 TTGCTAAAGAAAAGTGCAGAGAACAAAGCGATCTAGAACAGAGACGCTTAAG 196
DB 509 GTGCTAAAGAAAAGTGCAGAGAACAAAGCGATTTAGAACAGAGACGCTTAAG 568
QY 197 AAAAGTGCAGAGAACAAAGCGATTTAGAACAGAGACGCTTAAGAAAAGTTGC 256
DB 569 AAAAGTGCAGAGAACAAAGCGATTTAGAACAGAGACGCTTAAGAAAAGTTGC 628
QY 257 AAGAACAAAGAAAGCGATTTAGAACAGATAGACTTCTAAAGAAAAGTTACAGAGAC 316
DB 629 AAGAACAAAGAAAGCGATTTAGAACAGAGACTTCTAAAGAAAAGTTACAGAGAC 688
QY 317 AAAGCATTTAGAACAGAGAGACGCTTAAGAAAAGTTGCAGAGAACAAAGCGATT 376
DB 689 AAAGCATTTAGAACAGAGAGACGCTTAAGAAAAGTTGCAGAGAACAAAGCGATT 748
QY 377 TAGAACAGAGAGACGCTTAAGAAAAGTTGCAGAGAACAAAGCGATTTAGAACAG 436
DB 749 TAGAACAGAGAGACGCTTAAGAAAAGTTGCAGAGAACAAAGCGATTTAGAACAG 808
QY 437 AGAGACTTGTAAAGAAAAGTTGCAGAGAACAAAGCGATTTAGAACAGAGAGACGTG 496
DB 809 AGAGACTTGTAAAGAAAAGTTGCAGAGAACAAAGCGATTTAGAACAGAGAGACGTG 868
QY 497 CTAAAGAAAAGTTGCAGAGAACAAAGCGATTTAGAACAGAGAGACGCTTAAGAAA 556
DB 869 CTAAAGAAAAGTTGCAGAGAACAAAGCGATTTAGAACAGAGAGACGCTTAAGAAA 928
QY 557 AGTTGCAGAGAACAAAGCGATTTA 582
DB 929 GGTTCAGAGAACAAAGCGATTTA 954
```

```
RESULT 13
US-10-668-749A-1
; Sequence 1, Application US/10668749A
; Publication No. US20040110205A1
; GENERAL INFORMATION:
; APPLICANT: Agilent Technologies
; TITLE OF INVENTION: Methods and Systems for Nanopore Data Analysis
; FILE REFERENCE: 50112-1580
; CURRENT APPLICATION NUMBER: US/10/668,749A
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
OTHER INFORMATION: synthetic construct
US-10-668-749A-1
```

```
Query Match 15.7%; Score 233.4; DB 7; Length 1300;
Best Local Similarity 48.7%; Pred. No. 2.8e-23;
Matches 633; Conservative 0; Mismatches 666; Indels 0; Gaps 0;
```

```
QY 91 AAAGAAAAGTTACAGAGACGCAAAAGCGATTTAGAACAGAGAGACTTCTAAAGAAAAG 150
DB 1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60
QY 151 TTGCAAGAACAAAGCGATCTAGAACAGAGACGCTTAAGAAAAGTTGCAGAGAA 210
DB 61 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 120
QY 211 CAACAAAGCGATTTAGAACAGAGAGAGCTGCTTAAGAAAAGTTGCAGAGAACAAAGC 270
DB 121 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 180
QY 271 GATTTAGAACAGATAGACTTGTAAAGAAAAGTTACAGAGACGCAAGCGATTTAGAA 330
DB 181 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 240
QY 331 CAGAGAGACGCTTAAGAAAAGTTGCAGAGAACAAAGCGATTTAGAACAGAGAGA 390
DB 241 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 300
QY 391 CGTGTAAAGAAAAGTTCCAGAGAACAAAGCGATTTAGAACAGAGAGACTTGTAAA 450
DB 301 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 360
QY 451 GAAAAGTTGCAGAGAACAAAGCGATTTAGAACAGAGAGACGCTTAAGAAAAGTTG 510
DB 361 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 420
QY 511 CAGAGAACAAAGCGATTTAGAACAGAGAGACGCTTAAGAAAAGTTGCAGAGAACAA 570
DB 421 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 480
QY 571 CAAGCGATTTAGAACAGAGAGACGCTTAAGAAAAGTTGCAGAGAGACGAGAGAT 630
DB 481 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 540
QY 631 TTAGAACAAAGAGAGCGATTCGAAAAAATTTAGAAAGAAAAGGAAACATGAGAT 690
DB 541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 600
QY 691 ATATTAGAGAGATTTATATGCTGTTAGAAATACAGCTATAGAACTTCATCAGAA 750
DB 601 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 660
QY 751 AATGAACGTGATATTATATACCATCATCTTTTACCTCAGAGACAGAGGGAAT 810
DB 661 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 720
QY 811 AGTAGAGATTCAGAGAAATATCTATTAATAGAAAAAACAATAGAGAAATCTATTACA 870
DB 721 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 780
QY 871 AATGTTGAAGAGCAGAGGATATACATAGAGACATCTTGAGAAAAGAGAGATGTTCA 930
DB 781 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 840
QY 931 ATTAACCGAGAACAAAGAGATTAATCTGTGATACATCAAAATCATATTAGAGACA 990
DB 841 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 900
QY 991 GTAATATTCTGATGTTAATGATTTTCAATAGAGATAGAGATGAGATTAAGTGTCT 1050
DB 901 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 960
QY 1051 GAATATGACGATTCATTATATAGATGAGAGAGATGATGAAGACTTAGACGAATTTAAG 1110
```



```
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 461
; LENGTH: 6292
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-461

Query Match      10.2%; Score 151.2; DB 7; Length 6292;
Best Local Similarity 46.2%; Pred. No. 9, 2e-12;
Matches 619; Conservative 0; Mismatches 708; Indels 13; Gaps 3;

QY 40 AAAAAAAAAAGTTCAGACAGACAAACCGCTTTAGAACAGATGACTTCTTAAGAAAAG 99
   |||
DB 4548 AATATAAAAAAATTAATTAACCAAAAAAAATATACGAAACCGAAAAAATCATTA 4489
   |||

QY 100 TTACAAGACGACAAAGCGATTAGAACAGAGAGACTGCTAAAGAAAAGTTGCAGAA 159
   |||
DB 4488 CATAAAAACGATTAATAAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTTAAAA 4429
   |||

QY 160 CACAAAGCGATCTAGAACAGAGAGACGCTGCTAAAGAAAAGTTGCAGAACAGCAAGC 219
   |||
DB 4428 AACAATAATTAATTAACCAAAAAAAATTTGCAAAAAAATTAATTAATAAAAAAATT 4369
   |||

QY 220 GATTTAGAACAGAGAGACGCTGCTAAAGAAAAGTTGCAGAACAGCAAGCAATTTAGAA 279
   |||
DB 4368 TTTAATTAATTAATTAATAAAAAAATTAATAAAAAAATTAATAAAAAAATTAACAAAA 4309
   |||

QY 280 CAGATAGACTGCTTAAGAAAAGTTACAGAGACGACAAAGCGATTAGAAC---AAGA 335
   |||
DB 4308 AAAAAAATTTTACAAAAATTAATTAATAAAAAAATTTAAAAAATTTAAAAAATTTAAAA 4249
   |||

QY 336 GAGACGCTGCTAAAGAAAAGTTGCAGAACAGCAAGCGATTAGAACAGAGAGAGCGTGC 395
   |||
DB 4248 AAAAAATTTTAATTAATAAAAAAATTAATAAAAAAATTAATAAAAAAATTTAAAAAATA 4189
   |||

QY 396 TAAAGAAAAGTTGCAGAACAGCAAGCGATTAGAACAGAGAGAGCGTGCCTTAAGAAA 455
   |||
DB 4188 CAAAAAATTTTAATAAAAAAATTAATAAAAAAATTTTAATAAAAAAATTTTAATAAAAA 4129
   |||

QY 456 GTTGCAAGAACAGCAAGCGATTAGAACAGAGAGAGCGTGCCTTAAGAAAAGTTGCAGAA 515
   |||
DB 4128 CATTAATAAAAAAATTTTAATAATAAAAAAATTTAAAAAATTTAAACCAATTAATAAAA 4069
   |||

QY 516 ACAACAAAGCGATTAGAAC---AAGAGAGAGCGTCTAAAGAAAAGTTGCAGAACAAAC 571
   |||
DB 4068 AATATAAAAAATTAACGAACCGAAAAAATTTTAATAAAAAAATTTTAATAAAAAAATA 4009
   |||

QY 572 AAAAGCATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAGAGCAGCAAGAGATT 631
   |||
DB 4008 AAAAAAATTAATAAAAAAATTAATAAAAAAATTTTAATAAAAAAATTTTAATAAAAAAATA 3949
   |||

QY 632 TAGAACAAAGAGCGCTGATAGCAAAAAAATTTAGAAAAGAAAAGAAACATGAGATA 691
   |||
DB 3948 AAAAAATTAACAAATTAATAAAAAAATTTAAATTAACCAAAAAAATTTAAAAAATAAAC 3889
   |||

QY 692 TATTAGCAGAGATTATATATGCTGTTAGAAAATACAGCTATAGAACTTCATCAGAAA 751
   |||
DB 3888 TAAAAAATAAAAAAGATTAATAAAAAAATTTAATAATAAAAAAATAAAAAAATAACA 3829
   |||

QY 752 ATGAAAGCTGATATTATATACCATCATCTCTTTACTCAGGACACAGAGGGAATA 811
   |||
DB 3828 AACCAAAATTAATAAAAAAATTAATAAAAAAATTTTAATAAAAAAATTTTAATAAAAAAATA 3769
   |||

QY 812 GTAGAGATTCCAGAGAAATATCTATAATAGAAAAAATAATGAGAACTATTAACAACA 871
   |||
DB 3768 CAAAAAATAATTAATAAAAAAATTAATAAAAAAATTTAATAATAAAAAAATAAAAAAATA 3709
   |||

QY 872 ATGTTGAGAGACGAGAGATATATCATTAAGACATCTTGAGAGAAAAGAAAGATGTTCAA 931
   |||
```

```
DB 3708 AACTTACAAAAAAGCAAAAAAATATTTTAAATTAATAAAAAAATAAATAA 3649
QY 932 TAAACCGAACAAAAAGATTAATCTGCTGACATACAAATAATCATACATTAGAGACAG 991
   |||
DB 3648 CAAAAAATAAAAAAATTAATAAAAAAATTTTAAATTAACGAAAAAATAAAAAAATA 3589
   |||

QY 992 TAAATATTTCTGATGTTAATGATTTTCAATTAAGTAAATGATGAGATGAATAAGTGCT- 1050
   |||
DB 3588 AAAAAATTAACAGAACAAAAATTAATAAAAAAATTTTAATAAAAAAATAAAAAAATAATTT 3529
   |||

QY 1051 ----GAATATGACGATTCATTATATGATGAGAGAGAGATGATGAAGCTTAGACGAATT 1106
   |||
DB 3528 TTAATAAAAAAATAAAAAAATTAATAAAAAAATTAATAAAAAAATAAAAAAATAATTTACC 3469
   |||

QY 1107 TAGGCTATTGTGCAATATGACAAATTTCCAGATGAGAGAGAAACATAGCAATTTTAAGA 1166
   |||
DB 3468 AAAAAAATAAAAAAATTAATAAAAAAATTTTAAATTAATAAAAAAATAAAAAAATAACGA 3409
   |||

QY 1167 ACTAGAAATTTGATAGAGAAAAATGAATAATTTAGATGATTTAGATGAAGAAATAGAAA 1226
   |||
DB 3408 AATTAACGATTAATTAATAAAAAAATTTTACAAAAAATTAATAAAAAAATAAAAAAATA 3349
   |||

QY 1227 ATCATGAGAAATTTCTGAGAGAAAAATTAATAAAAAAAGAAATTAATAAAAAAATAACAA 1286
   |||
DB 3348 AAAAAATTTTAATAAAAAAATTAATAAAAAAATTTTAATAAAAAAATTTTAATAAAAAAATA 3289
   |||

QY 1287 GATTAATTAATTTTAAACCAATGATTAATAAGTTTGATGATGAGCATATTAAAAAATAATTA 1346
   |||
DB 3288 GAAAAAATAATTTCTCTCTTAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA 3229
   |||

QY 1347 AAATGATAGAGAGTTAATA 1366
   |||
DB 3228 TAAATTAATTAATTTCAAAA 3209
   |||
```

Search completed: December 7, 2005, 06:33:43
Job time : 1337 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 23:17:43 ; Search time 296 Seconds

(without alignments)
8899.819 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 1482

Sequence: 1 CAGGACACACAAAGCGATCT.....AATATTTCAGAACTAATA 1482

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PCUS_COMB.seq:*

7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*

8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*

9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1482	100.0	1482	US-08-098-327E-41	Sequence 41, Appl
2	1482	100.0	1482	US-08-462-625-41	Sequence 41, Appl
3	1480.4	99.9	1482	US-08-098-327E-45	Sequence 45, Appl
4	1480.4	99.9	1482	US-08-462-625-45	Sequence 45, Appl
5	1446	97.6	1493	US-08-098-327E-38	Sequence 38, Appl
6	1446	97.6	1493	US-08-462-625-38	Sequence 38, Appl
7	950	64.1	950	US-08-098-327E-32	Sequence 32, Appl
8	950	64.1	950	US-08-462-625-32	Sequence 32, Appl
9	853.2	57.6	1320	US-08-257-073-15	Sequence 15, Appl
10	472.8	31.9	988	US-08-098-327E-34	Sequence 34, Appl
11	472.8	31.9	988	US-08-462-625-34	Sequence 34, Appl
12	470.8	31.8	954	US-08-098-327E-37	Sequence 37, Appl
13	470.8	31.8	954	US-08-462-625-37	Sequence 37, Appl
14	129.8	8.2	1039	US-08-902-540-1280	Sequence 1280, Ap
15	120.8	8.2	612	US-09-902-540-1357	Sequence 1357, Ap
16	116.2	7.8	3095	5231168-1	Patent No. 5231168
17	111.6	7.5	1827	US-09-270-767-1308	Sequence 1308, Ap
18	111.6	7.5	1827	US-09-270-767-16590	Sequence 16590, A
19	110.2	7.4	7218	US-08-232-463-14	Sequence 14, Appl
20	101.2	6.8	1956	US-08-559-896B-1	Sequence 1, Appl
21	101.2	6.8	1956	US-09-351-794A-1	Sequence 1, Appl
22	95	6.4	187169	US-09-949-016-12775	Sequence 12775, A
23	95	6.4	191569	US-09-949-016-15940	Sequence 15940, A
24	94.8	6.4	614	US-09-902-540-1318	Sequence 1318, Ap

25	93.6	6.3	5361	3	US-08-973-462-2	Sequence 2, Appl
26	93.6	6.3	6152	3	US-08-973-462-1	Sequence 1, Appl
27	91.6	6.2	29927	3	US-09-949-016-11814	Sequence 11814, A
28	91.6	6.2	29927	3	US-09-949-016-17474	Sequence 17474, A
29	91.6	6.2	29927	3	US-09-949-016-17475	Sequence 17475, A
30	91.6	6.2	37802	3	US-09-949-016-12639	Sequence 12639, A
31	89.6	6.0	1664976	3	US-08-916-421B-1	Sequence 1, Appl
32	89.6	6.0	1664976	3	US-09-692-570-1	Sequence 1, Appl
33	88.4	6.0	5340	3	US-09-627-122-21	Sequence 21, Ap
34	87	5.9	639	3	US-09-248-796A-9722	Sequence 9722, Ap
35	86.8	5.9	4766	6	PCT-US93-07261-10	Sequence 10, Appl
36	86.4	5.8	601	3	US-09-949-016-203127	Sequence 203127,
37	86.4	5.8	601	3	US-09-949-016-203169	Sequence 203169,
38	85	5.7	147382	3	US-09-949-016-14624	Sequence 14624, A
39	83.2	5.6	1665	3	US-09-461-697-184	Sequence 184, Appl
40	82.8	5.6	19124	2	US-08-487-826B-113	Sequence 13, Appl
41	81.4	5.5	1685	3	US-10-104-047-1957	Sequence 1957, Ap
42	81.4	5.5	2223	2	US-08-257-073-4	Sequence 4, Appl
43	81.2	5.5	696	3	US-09-461-697-193	Sequence 193, App
44	81.2	5.5	699	3	US-09-461-697-191	Sequence 191, App
45	81.2	5.5	717	3	US-09-461-697-189	Sequence 189, App

ALIGNMENTS

RESULT 1
US-08-098-327E-41
; Sequence 41, Application US/08098327E
; Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/33884
PUBLICATION DATE: 20-AUG-1992

US-08-098-337E-41

Query Match 100.0%; Score 1482; DB 3; Length 1482;
Best Local Similarity 100.0%; Pred. No. 1,88-246;
Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CAAAGAACAAAGCGATCTAGAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGACAA 60
DB 1 CAAAGAACAAAGCGATCTAGAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGACAA 60
QY 61 CAAAGCGATTTGAAACAAGATAGACTTGTAAAGAAAGTTGCAAGAGAGAGCGAT 120
DB 61 CAAAGCGATTTGAAACAAGATAGACTTGTAAAGAAAGTTGCAAGAGAGAGCGAT 120
QY 121 TTAGAACAAAGAGAGAGCTTGTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
DB 121 TTAGAACAAAGAGAGAGCTTGTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
QY 181 GAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGT 240
DB 181 GAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGT 240
QY 241 GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGTAAAGAA 300
DB 241 GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGTAAAGAA 300
QY 301 AAGTTACAAAGAGAGAGAGCTTTAGAACAAAGAGAGAGCGTCTAAAGAAAGTTGCA 360
DB 301 AAGTTACAAAGAGAGAGAGCTTTAGAACAAAGAGAGAGCGTCTAAAGAAAGTTGCA 360
QY 361 GAAACAAACAAAGCGATTTAGAACAAAGAGAGAGCGTCTAAAGAAAGTTGCAAGACAA 420
DB 361 GAAACAAACAAAGCGATTTAGAACAAAGAGAGAGCGTCTAAAGAAAGTTGCAAGACAA 420
QY 421 ACGGATTTAGAACAAAGAGAGCTTGTAAAGAAAGTTGCAAGAACAAAGCGATTTA 480
DB 421 ACGGATTTAGAACAAAGAGAGCTTGTAAAGAAAGTTGCAAGAACAAAGCGATTTA 480
QY 481 GAAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAG 540
DB 481 GAAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAG 540
QY 541 AGACGCTGTTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCGTCT 600
DB 541 AGACGCTGTTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCGTCT 600
QY 601 AAAAGAAAGTTGCAAGAGCAAGCAAGATTTAGAACAAAGAGAGCGTCTATCGAAAAA 660
DB 601 AAAAGAAAGTTGCAAGAGCAAGCAAGATTTAGAACAAAGAGAGCGTCTATCGAAAAA 660
QY 661 AATTGAAAGAAAGAAAGCAATGAGATATTTAGAGAGAGATTTATATGTCGTTTA 720
DB 661 AATTGAAAGAAAGAAAGCAATGAGATATTTAGAGAGAGATTTATATGTCGTTTA 720
QY 721 GAAATATCCAGCTATATAGACTTCATCAGAAATGAAAGTATATATATACCACTCAA 780
DB 721 GAAATATCCAGCTATATAGACTTCATCAGAAATGAAAGTATATATATACCACTCAA 780
QY 781 TCTTTCTTTTACCTCAGAGCAAGAGAGATATGAGAGATTTCCAGAGAAATATCTATAA 840
DB 781 TCTTTCTTTTACCTCAGAGCAAGAGAGATATGAGAGATTTCCAGAGAAATATCTATAA 840
QY 841 GAAAAAACAATGAGATCTATATTAACAACAAATGTTAGAGAGAGAGATTAACATAA 900
DB 841 GAAAAAACAATGAGATCTATATTAACAACAAATGTTAGAGAGAGAGATTAACATAA 900
QY 901 GGCATCTTTGAGAGAAAGAAAGATGTTCAATTAACAAAGCAAGAAAGAAAGATTAATCT 960
DB 901 GGCATCTTTGAGAGAAAGAAAGATGTTCAATTAACAAAGCAAGAAAGAAAGATTAATCT 960

```

```

QY 961 GCTGACATCAAAATCATATACATTAGAGACAGTAATAATTTCTGATGTTAATGATTTTCAA 1020
DB 961 GCTGACATCAAAATCATATACATTAGAGACAGTAATAATTTCTGATGTTAATGATTTTCAA 1020
QY 1021 ATTAAGTAAGTATGAGATGAAATATAGTCTGATATATGACATTTCAATTAATGATGAAGA 1080
DB 1021 ATTAAGTAAGTATGAGATGAAATATAGTCTGATATATGACATTTCAATTAATGATGAAGA 1080
QY 1081 GAAATGATGAGAGCTTATAGAGCAATTTAAGCTTATGTCATATGACATTTTCCAGAT 1140
DB 1081 GAAATGATGAGAGCTTATAGAGCAATTTAAGCTTATGTCATATGACATTTTCCAGAT 1140
QY 1141 GAGAAAAACATAGCAATTTATTAAGAACTAGACAGATTTGATAGAGAAAAATGAAATTTA 1200
DB 1141 GAGAAAAACATAGCAATTTATTAAGAACTAGACAGATTTGATAGAGAAAAATGAAATTTA 1200
QY 1201 GATGATTTAGTGAAGATGAAATATCATCAGAGATTTATCTGAAGAAAAATTA 1260
DB 1201 GATGATTTAGTGAAGATGAAATATCATCAGAGATTTATCTGAAGAAAAATTA 1260
QY 1261 AAAGAAAGAAATATGAAAAAACAAGGATATATTTTAAACCAATGATTAAGTTG 1320
DB 1261 AAAGAAAGAAATATGAAAAAACAAGGATATATTTTAAACCAATGATTAAGTTG 1320
QY 1321 TATGATGACATATTTAAATAATATTAATAATGATAGACAGTTAATAGAAAAAGAAAA 1380
DB 1321 TATGATGACATATTTAAATAATATTAATAATGATAGACAGTTAATAGAAAAAGAAAA 1380
QY 1381 TTCAATTAATCATATTTTCAATATTTTGAAGAGCAATGAAATTTTACAGATCGTGAT 1440
DB 1381 TTCAATTAATCATATTTTCAATATTTTGAAGAGCAATGAAATTTTACAGATCGTGAT 1440
QY 1441 GAGTTATCTGAAGATATATTAATTTTATGAACTATTA 1482
DB 1441 GAGTTATCTGAAGATATATTAATTTTATGAACTATTA 1482

```

RESULT 2

US-08-462-625-41

/ Sequence 41, Application US/08462625

/ Patent No. 6319502

/ GENERAL INFORMATION:

/ APPLICANT: GUERIN-MARCHAND, Claudine

/ APPLICANT: DRIJLHE, Pierre

/ TITLE OF INVENTION: HEPATIC STAGES OF P. PALCIPARUM BEARING EPITOPES CAPABLE

/ TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES

/ NUMBER OF SEQUENCES: 46

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Burns, Doane, Swecker & Mathis

/ STREET: P. O. Box 1404

/ CITY: Alexandria

/ STATE: Virginia

/ COUNTRY: United States

/ ZIP: 22113-1404

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patentin Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/462,625

/ FILING DATE: 05-JUN-1995

/ CLASSIFICATION: 424

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US 08/098,327

/ FILING DATE: 24-NOV-1993

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: FR 91 01286

/ FILING DATE: 05-FEB-1991

/ ATTORNEY/AGENT INFORMATION:

/ NAME: McGowan, Malcolm K.

/ REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-078
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1482 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1482
 PUBLICATION INFORMATION:
 DOCUMENT NUMBER: WO 92/13884
 PUBLICATION DATE: 20-AUG-1992
 US-08-462-625-41

Query Match 100.0%; Score 1482; DB 3; Length 1482;
 Best Local Similarity 100.0%; Pred. No. 1,8e-246;
 Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGAACAACAAGCGATCTAGAACAGAGAGAGTGTCTAAAGAAAAGTTGCAAGAACAA 60
 DB 1 CAAGAACAACAAGCGATCTAGAACAGAGAGAGTGTCTAAAGAAAAGTTGCAAGAACAA 60
 QY 61 CAAAGCGATTTAGAACAGATAGACTTCTAAAGAAAAGTTGCAAGAGCAGCAAGCGAT 120
 DB 61 CAAAGCGATTTAGAACAGATAGACTTCTAAAGAAAAGTTGCAAGAGCAGCAAGCGAT 120
 QY 121 TTAGAACAGAGAGACTTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
 DB 121 TTAGAACAGAGAGACTTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
 QY 181 GAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCT 240
 DB 181 GAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCT 240
 QY 241 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATAGACTTGTCTAAAGAA 300
 DB 241 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATAGACTTGTCTAAAGAA 300
 QY 301 AAGTTACAGAGAGCAGCAAGCGATTTAGAACAGAGAGAGCTGTCTAAAGAAAAGTTGCA 360
 DB 301 AAGTTACAGAGAGCAGCAAGCGATTTAGAACAGAGAGAGCTGTCTAAAGAAAAGTTGCA 360
 QY 361 GAAACAACAAGCGATTTAGAACAGAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAA 420
 DB 361 GAAACAACAAGCGATTTAGAACAGAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAA 420
 QY 421 AGCGATTTAGAACAGAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
 DB 421 AGCGATTTAGAACAGAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
 QY 481 GAACAAGAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 540
 DB 481 GAACAAGAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 540
 QY 541 AGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCTGTCT 600
 DB 541 AGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCTGTCT 600
 QY 601 AAAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGAGCTGATACGAAAAA 660
 DB 601 AAAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGAGCTGATACGAAAAA 660
 QY 661 AATTAGAAAAGAAAAGAGAACTAGAGATATATTAGCAGAGAGATTTATATGTCGTTTA 720
 DB 661 AATTAGAAAAGAAAAGAGAACTAGAGATATATTAGCAGAGAGATTTATATGTCGTTTA 720
 QY 721 GAATATCAAGCTATAGAACTTCCATCAAAAAATGAACCTGATATTAATATCCACATCAA 780
 DB 721 GAATATCAAGCTATAGAACTTCCATCAAAAAATGAACCTGATATTAATATCCACATCAA 780

DB 721 GAATATCAAGCTATAGAACTTCCATCAAAAAATGAACCTGATATTAATATCCACATCAA 780
 QY 781 TCTTCTTTTACCTAGAGCAACAGAGGAATATGATGATTTCCAGAAAATCTATATATA 840
 DB 781 TCTTCTTTTACCTAGAGCAACAGAGGAATATGATGATTTCCAGAAAATCTATATATA 840
 QY 841 GAAAAAACAATAGAGATCTATTAACAACAATGTTGAAGAGAGAGATATACATATA 900
 DB 841 GAAAAAACAATAGAGATCTATTAACAACAATGTTGAAGAGAGAGATATACATATA 900
 QY 901 GGCATCTTGAAGAAAAGAAAGTGTCTCAATATAACAGAACAAAAAGAGTAAATCT 960
 DB 901 GGCATCTTGAAGAAAAGAAAGTGTCTCAATATAACAGAACAAAAAGAGTAAATCT 960
 QY 961 GCTGACATTAACAATATCATATTAGAGACAGTAAATTTTCTGATGTTATGATTTTCA 1020
 DB 961 GCTGACATTAACAATATCATATTAGAGACAGTAAATTTTCTGATGTTATGATTTTCA 1020
 QY 1021 ATTAAGTATGATAGAGATGAAATTAAGTGTCAATATGACGATTCATTAATAGATGAAGA 1080
 DB 1021 ATTAAGTATGATAGAGATGAAATTAAGTGTCAATATGACGATTCATTAATAGATGAAGA 1080
 QY 1081 GAAGATGATGAGACCTTAGACGAATTTAAGCTATTTGTGCAATATGACAAATTTTCAAGAT 1140
 DB 1081 GAAGATGATGAGACCTTAGACGAATTTAAGCTATTTGTGCAATATGACAAATTTTCAAGAT 1140
 QY 1141 GAAAGAAAACATAGCAATTTTAAAGAACTTGAAAGATTTGATAGAGAAAATGAAAATTTA 1200
 DB 1141 GAAAGAAAACATAGCAATTTTAAAGAACTTGAAAGATTTGATAGAGAAAATGAAAATTTA 1200
 QY 1201 GATGATTTAGATAGAGAAATGAGAAAATCATCAGAGAAATTAATCTGAGAGAAAATTA 1260
 DB 1201 GATGATTTAGATAGAGAAATGAGAAAATCATCAGAGAAATTAATCTGAGAGAAAATTA 1260
 QY 1261 AAAGAAAAGAAATATGAGAAAACAAAGATATATTTTAAACCAATATGATATAAAGTTTG 1320
 DB 1261 AAAGAAAAGAAATATGAGAAAACAAAGATATATTTTAAACCAATATGATATAAAGTTTG 1320
 QY 1321 TATGATGAGATATTAATAAATATTAATAATGATATGACAGTTAATTAAGAAAAGGAAAA 1380
 DB 1321 TATGATGAGATATTAATAAATATTAATAATGATATGACAGTTAATTAAGAAAAGGAAAA 1380
 QY 1381 TTTCTTAAATCATTTGTTTCTATATTTGACGAGACATATAATTTTACATGCTGAT 1440
 DB 1381 TTTCTTAAATCATTTGTTTCTATATTTGACGAGACATATAATTTTACATGCTGAT 1440
 QY 1441 GAGTTATCTGAGATATTAATAATATTTATATGAACTATATA 1482
 DB 1441 GAGTTATCTGAGATATTAATAATATTTATATGAACTATATA 1482

RESULT 3
 US-08-098-327E-45
 ; Sequence 45. Application US/08098327E
 ; Patent No. 6270771
 ; GENERAL INFORMATION:
 ; APPLICANT: GUERIN-MARCHAND, Claudine
 ; APPLICANT: DRULHE, Pierre
 ; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
 ; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
 ; OF STIMULATING THE T LYMPHOCYTES
 ; NUMBER OF SEQUENCES: 46
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis
 ; STREET: P.O. Box 1404
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: United States
 ; ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-45

Query Match 99.9%; Score 1480.4; DB 3; Length 1482;
Best Local Similarity 99.9%; Pred. No. 3.5e-246;
Matches 1481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAGAACAAACGATCTTAAACAAAGAGACGCTCTAAAGAAAAGTTGCAAGACAA 60
DB 1 CAAGAACAAACGATCTTAAACAAAGAGACGCTCTAAAGAAAAGTTGCAAGACAA 60
QY 61 CAAAGGATTTAGAACAAATAGACTTGTAAAGAAAAGTTACAGAGCAAGACGAT 120
DB 61 CAAAGGATTTAGAACAAATAGACTTGTAAAGAAAAGTTACAGAGCAAGACGAT 120
QY 121 TTAGAACAAAGAGACTTCTAAAGAAAAGTTGCAAGACAAACGATCTAGAACAA 180
DB 121 TTAGAACAAAGAGACTTCTAAAGAAAAGTTGCAAGACAAACGATCTAGAACAA 180
QY 181 GAGAGACGCTTAAAGAAAAGTTGCAAGACAAACGATTTAGAACAAAGAGACGT 240
DB 181 GAGAGACGCTTAAAGAAAAGTTGCAAGACAAACGATTTAGAACAAAGAGACGT 240
QY 241 GCTAAAGAAAAGTTGCAAGACAAACGATTTAGAACAAAGTCTCTAAAGAA 300
DB 241 GCTAAAGAAAAGTTGCAAGACAAACGATTTAGAACAAAGTCTCTAAAGAA 300
QY 301 AAGTTCAAGAGACGACAAACGATTTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCAA 360
DB 301 AAGTTCAAGAGACGACAAACGATTTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCAA 360
QY 361 GAAACAACAAAGCGATTTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCAAGACAA 420
DB 361 GAAACAACAAAGCGATTTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCAAGACAA 420
QY 421 AGCGATTTAGAACAAAGAGACCTTGTAAAGAAAAGTTGCAAGACAAACGATTTA 480
DB 421 AGCGATTTAGAACAAAGAGACCTTGTAAAGAAAAGTTGCAAGACAAACGATTTA 480
QY 481 GAAACAAGAGACGCTCTAAAGAAAAGTTGCAAGACAAACGATTTAGAACAAAGAG 540
DB 481 GAAACAAGAGACGCTCTAAAGAAAAGTTGCAAGACAAACGATTTAGAACAAAGAG 540
QY 541 AGACGCTCTAAAGAAAAGTTGCAAGACAAACGATTTAGAACAAAGAGACGCTCT 600
DB 541 AGACGCTCTAAAGAAAAGTTGCAAGACAAACGATTTAGAACAAAGAGACGCTCT 600

QY 601 AAAGAAAAGTTGCAAGACGACAAAGAGATTTAGAACAAAGAAAGCTGATACGAAAAA 660
DB 601 AAAGAAAAGTTGCAAGACGACAAAGAGATTTAGAACAAAGAAAGCTGATACGAAAAA 660
QY 661 AATTTAGAAAAGAAAAGAAACATGAGATATATTTAGCAGAGATTTATATGCTGTTA 720
DB 661 AATTTAGAAAAGAAAAGAAACATGAGATATATTTAGCAGAGATTTATATGCTGTTA 720
QY 721 GAAATACGCTATATGAACTTCCATCAGAAAATGAGTGAATTTATATACCATCAA 780
DB 721 GAAATACGCTATATGAACTTCCATCAGAAAATGAGTGAATTTATATACCATCAA 780
QY 781 TCTTCTTACCTCAGACCAAGAGGAAATGAGATTTCCAGAAAATATCTATTAATA 840
DB 781 TCTTCTTACCTCAGACCAAGAGGAAATGAGATTTCCAGAAAATATCTATTAATA 840
QY 841 GAAAAACAAATAGAGATCTATTTACAAACAAATGTTGAAAGACGAAAGGATATACATAA 900
DB 841 GAAAAACAAATAGAGATCTATTTACAAACAAATGTTGAAAGACGAAAGGATATACATAA 900
QY 901 GGAACATCTTGAAGAAAAGATGCTCAATTAACAGAAACAAAGAAAGATTAATCT 960
DB 901 GGAACATCTTGAAGAAAAGATGCTCAATTAACAGAAACAAAGAAAGATTAATCT 960
QY 961 GCTGACATACAAAATCATACATTTAGAGACAGTAAATATTTCTGATTTAATGATTTTCAA 1020
DB 961 GCTGACATACAAAATCATACATTTAGAGACAGTAAATATTTCTGATTTAATGATTTTCAA 1020
QY 1021 ATATGATATATGAGAGATGAATAATGCTGTAATATGCAATTCATTAATAGTAAAGA 1080
DB 1021 ATATGATATATGAGAGATGAATAATGCTGTAATATGCAATTCATTAATAGTAAAGA 1080
QY 1081 GAAAGTATGAAAGCTTATGAGAAATTTAAGCTATTTGCAATATAGCAATTTCCAGAT 1140
DB 1081 GAAAGTATGAAAGCTTATGAGAAATTTAAGCTATTTGCAATATAGCAATTTCCAGAT 1140
QY 1141 GAAAGAAACATATGAAATTTATTAAGAACTAGAGATTTGATAGAAAAATGAAATTTA 1200
DB 1141 GAAAGAAACATATGAAATTTATTAAGAACTAGAGATTTGATAGAAAAATGAAATTTA 1200
QY 1201 GATGATTTAGATGAAGAAATATGAAATATCATCAGAAATTTATCGAAGAAAAATTA 1260
DB 1201 GATGATTTAGATGAAGAAATATGAAATATCATCAGAAATTTATCGAAGAAAAATTA 1260
QY 1261 AAAGAAAAGAAATATGAAAGAAACAAAGATATATTTTAAACCAATGATTAAGTTG 1320
DB 1261 AAAGAAAAGAAATATGAAAGAAACAAAGATATATTTTAAACCAATGATTAAGTTG 1320
QY 1321 TATGATGACATATTAATAAATATTAATAATGATAAGAGTTAATAGAAAAAGAAAA 1380
DB 1321 TATGATGACATATTAATAAATATTAATAATGATAAGAGTTAATAGAAAAAGAAAA 1380
QY 1381 TTCATTAATATCTTTTCAATATATTTTGAACGAGCAATGAATTTTACAGATCGTGAT 1440
DB 1381 TTCATTAATATCTTTTCAATATATTTTGAACGAGCAATGAATTTTACAGATCGTGAT 1440
QY 1441 GAGTATATCGAAGATATTAATATTTTGAAGAACTATTA 1482
DB 1441 GAGTATATCGAAGATATTAATATTTTGAAGAACTATTA 1482

RESULT 4
US-08-462-625-45
Sequence 45, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPS CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burnb, Doane, Swecker & Mathis
;; STREET: P.O. Box 1404
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/462,625
;; FILING DATE: 05-JUN-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/098,327
;; FILING DATE: 24-NOV-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: FR 91 01286
;; FILING DATE: 05-FEB-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McGowan, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 010830-078
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-6620
;; INFORMATION FOR SEQ ID NO: 45:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1482 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..1482
;; PUBLICATION INFORMATION:
;; DOCUMENT NUMBER: WO 92/13884
;; PUBLICATION DATE: 20-AUG-1992
US-08-462-625-45
Query Match 99.9%; Score 1480.4; DB 3; Length 1482;
Beet Local Similarity 99.9%; Pred. No. 3.5e-246;
Matches 1481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAGAACAAAGCGATCTGAAACAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 60
DB 1 CAAGAACAAAGCGATCTGAAACAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 60
QY 61 CAAGCGATTGAAACAAAGATGAGCTTGCTAAAGAAAAGTTGCAAGAGCAGCAAGCGAT 120
DB 61 CAAGCGATTGAAACAAAGATGAGCTTGCTAAAGAAAAGTTGCAAGAGCAGCAAGCGAT 120
QY 121 TTAGAACAAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
DB 121 TTAGAACAAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
QY 181 GAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCT 240
DB 181 GAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCT 240
QY 241 GCTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATGAGCTTGCTAAAGAA 300
DB 241 GCTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATGAGCTTGCTAAAGAA 300
QY 301 AAGTTACAAGAGCAGCAAGCGATTTAGAACAAAGAGAGAGCGTCTAAAGAAAAGTTGCA 360
DB 301 AAGTTACAAGAGCAGCAAGCGATTTAGAACAAAGAGAGAGCGTCTAAAGAAAAGTTGCA 360
QY 361 GAACAAACAAGCGATTTAGAACAAAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 420

DB 361 GAACAAACAAGCGATTTAGAACAAAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 420
QY 421 AGCGATTTAGAACAAAGAGAGCTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
DB 421 AGCGATTTAGAACAAAGAGAGCTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
QY 481 GAACAAAGAGAGAGCTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 540
DB 481 GAACAAAGAGAGAGCTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 540
QY 541 AGAGTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGTCT 600
DB 541 AGAGTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGTCT 600
QY 601 AAAGAAAAGTTGCAAGAGCAGCAAGAGATTTGAAACAAAGAGCGTGTATGAAAAAA 660
DB 601 AAAGAAAAGTTGCAAGAGCAGCAAGAGATTTGAAACAAAGAGCGTGTATGAAAAAA 660
QY 661 AATTTAGAAAAGAAAAGAAACATGAGATATTTAGAGAGATTTATATGATGCTGTTA 720
DB 661 AATTTAGAAAAGAAAAGAAACATGAGATATTTAGAGAGATTTATATGATGCTGTTA 720
QY 721 GAATACACAGCTATGAACTTCCATCAGAAAATGAAAGTGTATTTATATATACACATCAA 780
DB 721 GAATACACAGCTATGAACTTCCATCAGAAAATGAAAGTGTATTTATATATACACATCAA 780
QY 781 TCTTCTTTACCTCAGAGCAACAGAGGAAATGAGAGATTTCCAGGAATATCTATATATA 840
DB 781 TCTTCTTTACCTCAGAGCAACAGAGGAAATGAGAGATTTCCAGGAATATCTATATATA 840
QY 841 GAAAAAACAAATGAGAAATTTATTAACAACAAAGTTGAAGAGAGAGGATATACATAAA 900
DB 841 GAAAAAACAAATGAGAAATTTATTAACAACAAAGTTGAAGAGAGAGGATATACATAAA 900
QY 901 GGCATCTTGAAGAAAAGAAAGATGTTCAATATAAACAGAACAAAGAAAGATTAATCT 960
DB 901 GGCATCTTGAAGAAAAGAAAGATGTTCAATATAAACAGAACAAAGAAAGATTAATCT 960
QY 961 GCTGACATCAAAAATCATCATTTAGAGACAGTAAATTTCTGATGTTAATGATTTTCAA 1020
DB 961 GCTGACATCAAAAATCATCATTTAGAGACAGTAAATTTCTGATGTTAATGATTTTCAA 1020
QY 1021 ATTAAGTAAGTATGAGATGAAATTAAGTCTGAATTAAGACATTAATAGATGAAGAA 1080
DB 1021 ATTAAGTAAGTATGAGATGAAATTAAGTCTGAATTAAGACATTAATAGATGAAGAA 1080
QY 1081 GAAGATGATGAAGCTTAGAGCAATTTAAGCTTATGTCGAATATGACAAATTTCCAAGAT 1140
DB 1081 GAAGATGATGAAGCTTAGAGCAATTTAAGCTTATGTCGAATATGACAAATTTCCAAGAT 1140
QY 1141 GAAGAAAACATAGGAATTTATTAAGAACTAGAAAGTTGATAGAGAAAATGAAAATTTA 1200
DB 1141 GAAGAAAACATAGGAATTTATTAAGAACTAGAAAGTTGATAGAGAAAATGAAAATTTA 1200
QY 1201 GATGATTTAGATGAAGAAATGAGAAATCATCAGAGAAATTTATCTGAGAAAATATAAA 1260
DB 1201 GATGATTTAGATGAAGAAATGAGAAATCATCAGAGAAATTTATCTGAGAAAATATAAA 1260
QY 1261 AAAGGAAAAGAAATTTGAAAAAAACAAAGATTAATTTTAAACCAATGATTAAGATTGG 1320
DB 1261 AAAGGAAAAGAAATTTGAAAAAAACAAAGATTAATTTTAAACCAATGATTAAGATTGG 1320
QY 1321 TATGATAGCATATTAATAAATATTAATAATGATTAAGCAGGTTAATAAGAAAAGAAAA 1380
DB 1321 TATGATAGCATATTAATAAATATTAATAATGATTAAGCAGGTTAATAAGAAAAGAAAA 1380
QY 1381 TTCAATAAATCATTTGTTTCATATTTTGAAGAGAGACAAAGAAATTTTACAGATCGTGGAT 1440
DB 1381 TTCAATAAATCATTTGTTTCATATTTTGAAGAGAGACAAATAAATTTTACAGATCGTGGAT 1440
QY 1441 GAGTTATCTGAGATATATACATAAATATTTATGAACTATATA 1482

Db 1441 GAGTATCTGAGATATACTAAATATTTTATGAACATATA 1482

RESULT 5
US-08-098-327E-38
Sequence 38, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-38

Query Match 97.6%; Score 1446; DB 3; Length 1493;
Best Local Similarity 99.8%; Pred. No. 2,9e-240;
Matches 1479; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 1 CAAGAACAACAAGCATCTAGAACAAAGAGACGCTTAAGAAAGTTGCAAGAACAA 60
Db 1 CAAGAACAACAAGCATCTAGAACAAAGAGACGCTTAAGAAAGTTGCAAGAACAA 60
QY 61 CAAGCGATTTAGAACAGATAGACTTGTAAAGAAAGTTACAAAGACGCAAGCGAT 120
Db 61 CAAGCGATTTAGAACAGATAGACTTGTAAAGAAAGTTACAAAGACGCAAGCGAT 120
QY 121 TTGAACAAGAGAGACTTGTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
Db 121 TTGAACAAGAGAGACTTGTCT -AAGAAAGTTGCAAGAACAAAGCGATCTAGAACAA 179
QY 181 GAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGT 240
Db 181 GAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGT 239
QY 241 GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATGACTTGTAAAGAA 300

Db 240 GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATGACTTGTAAAGAA 299
QY 301 AAGTTACAAAGAGCAAAAGCGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAA 360
Db 300 AAGTTACAAAGAGCAAAAGCGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAA 359
QY 361 GAACAAACAAGCGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAAGAACAA 420
Db 360 GAACAAACAAGCGATTTAGAACAAAGAGACGTCCT -AAGAAAGTTGCAAGAACAA 418
QY 421 AGCGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 480
Db 419 AGCGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 478
QY 481 GAACAGAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAG 540
Db 479 GAACAGAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAG 538
QY 541 AGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGTCGT 600
Db 539 AGACGTCGT -AAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGTCGT 597
QY 601 AAAGAAAGTTGCAAGAGACGCAAGAGATTTAGAACAAAGAGAGCTGATCGAAAAA 660
Db 598 AAAGAAAGTTGCAAGAGACGCAAGAGATTTAGAACAAAGAGAGCTGATCGAAAAA 657
QY 661 AATTAGAAAGAAAGAAAGAAATGAGATATATAGAGAGATTTATATGTCGTTA 720
Db 658 AATTAGAAAGAAAGAAAGAAATGAGATATATAGAGAGATTTATATGTCGTTA 717
QY 721 GAATATACCAAGCTATAGAACTTCATCAGAAAAATGAACTGATATATATACACATCAA 780
Db 718 GAATATACCAAGCTATAGAACTTCATCAGAAAAATGAACTGATATATATACACATCAA 777
QY 781 TCTTCTTTTAACTCAGAGAACAAAGAGGAAATAGAGATTCAGAGAAATATCTTAATA 840
Db 778 TCTTCTTTTAACTCAGAGAACAAAGAGGAAATAGAGATTCAGAGAAATATCTTAATA 837
QY 841 GAAAAAACAATATAGAAATCTATTACAAACAATGTTGAAGCAAGAGGATATACATAA 900
Db 838 GAAAAAACAATATAGAAATCTATTACAAACAATGTTGAAGCAAGAGGATATACATAA 897
QY 901 GCACATCTTGAAGAAAGAAAGATGTTCAATTAACCAAGAACAAAGAAATATAATCT 960
Db 898 GCACATCTTGAAGAAAGAAAGATGTTCAATTAACCAAGAACAAAGAAATATAATCT 957
QY 961 GCTGACATACAAATCATACATTAGAGACAGTAATATTTCTGATGTTAATGATTTTCAA 1020
Db 958 GCTGACATACAAATCATACATTAGAGACAGTAATATTTCTGATGTTAATGATTTTCAA 1017
QY 1021 AATAGATATGAGAGATGAAATTAAGTCTGAATATGACGATTCATTAATAGATGAAGA 1080
Db 1018 AATAGATATGAGAGATGAAATTAAGTCTGAATATGACGATTCATTAATAGATGAAGA 1077
QY 1081 GAAGATGATGAAGACTTGAAGCAATTTAAGCCTATATGCAATATGACAAATTTCCAAAG 1140
Db 1078 GAAGATGATGAAGACTTGAAGCAATTTAAGCCTATATGCAATATGACAAATTTCCAAAG 1137
QY 1141 GAAGAAACAATGGAATTTATTAAGAACTAGAAAGTTGATGAGAAAAATGAAATTTTA 1200
Db 1138 GAAGAAACAATGGAATTTATTAAGAACTAGAAAGTTGATGAGAAAAATGAAATTTTA 1197
QY 1201 GATGATTTAGAGAAAGAAATTAAGAAATCATCAGAGAAATTTCTGAAGAAAAATTA 1260
Db 1198 GATGATTTAGAGAAAGAAATTAAGAAATCATCAGAGAAATTTCTGAAGAAAAATTA 1257
QY 1261 AAAGAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATATGATTAAGTTTG 1320
Db 1258 AAAGAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATATGATTAAGTTTG 1317
QY 1321 TATGATAGCATATTAATAATATTAATAATGATTAAGCAGTTATTAAGAAAAAGCAAAA 1380

Db 1318 TATGATAGCATATTAATAAATATATAAATGATAGAGGTTAATAGAAAGAAAA 1377

Qy 1381 TTCATAAAATCATTTGTTTCATATATTTGACGAGACAAATGAAATTTTACAGATCGTGAT 1440

Db 1378 TTCATAAAATCATTTGTTTCATATATTTGACGAGACAAATGAAATTTTACAGATCGTGAT 1437

Qy 1441 GAGTATCTGAGATATATATATATTTTATGAACTATTA 1482

Db 1438 GAGTTATCTGAGATATATATATTTTATGAACTATTA 1479

RESULT 6

US-08-462-625-38

Sequence 38, Application US/08462625

Patent No. 6319502

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine

APPLICANT: DRULHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE

TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE

TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,625

FILING DATE: 05-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/098,327

FILING DATE: 24-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 91 01286

FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-078

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:

LENGTH: 1493 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992

US-08-462-625-38

Query Match 97.6%; Score 1446; DB 3; Length 1493;

Best Local Similarity 99.8%; Pred. No. 2.9e-240;

Matches 1479; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

Qy 1 CAAGAACAAAGGATCTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAACAA 60

Db 1 CAAGAACAAAGGATCTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAACAA 60

Qy 61 CAAGCGGTTAGAACAGATGAGCTTGCTAAAGAAAGTTAGCAAGAGCAGCAAGCGAT 120

Db 61 CAAGCGATTTAGAACAGATGAGCTTGCTAAAGAAAGTTAGCAAGAGCAGCAAGCGAT 120

Qy 121 TTGAAACAGAGAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATCTTAGAACAA 180

Db 121 TTGAAACAGAGAGACTTGTCT - AAGAAAGTTGCAAGAACAAAGCGATCTTAGAACAA 179

Qy 181 GAGAGAGCTGTAAAGAAAGTTGCAAGAACAAAGCATTTAGAACAGAGAGAGCT 240

Db 180 GAGAGAGCTGTCTAAAGAAAGTTGCAAGAACAAAGCATTTAGAACAGAGAGAGCT 239

Qy 241 GCTAAAGAAAGTTGCAAGAACAAAGCATTTAGAACAGATGAGCTTGTAAAGAA 300

Db 240 GCTAAAGAAAGTTGCAAGAACAAAGCATTTAGAACAGATGAGCTTGTAAAGAA 299

Qy 301 AAGTTACAGAGAGCAGCAAGCGATTTAGAACAGAGAGAGCTGTAAAGAAAGTTGCCAA 360

Db 300 AAGTTACAGAGAGCAGCAAGCGATTTAGAACAGAGAGAGCTGTAAAGAAAGTTGCCAA 359

Qy 361 GAACCAACAAAGCGATTTAGAACAGAGAGAGCTGTAAAGAAAGTTGCCAAAGAACAA 420

Db 360 GAACCAACAAAGCGATTTAGAACAGAGAGAGCTGT - AAGAAAGTTGCCAAAGAACAA 418

Qy 421 AGCGATTTAGAACAGAGAGACTTGTCTAAAGAAAGTTGCCAAAGAACAAAGCGATTTA 480

Db 419 AGCGATTTAGAACAGAGAGACTTGTCTAAAGAAAGTTGCCAAAGAACAAAGCGATTTA 478

Qy 481 GAACAGAGAGAGCTGTCTAAAGAAAGTTGCCAAAGAACAAAGCGATTTAGAACAGAG 540

Db 479 GAACAGAGAGAGCTGTCTAAAGAAAGTTGCCAAAGAACAAAGCGATTTAGAACAGAG 538

Qy 541 AGAGCTGTAAAGAAAGTTGCCAAAGAACAAAGCATTTAGAACAGAGAGAGCTGTCT 600

Db 539 AGAGCTGT - AAGAAAGTTGCCAAAGAACAAAGCATTTAGAACAGAGAGAGCTGTCT 597

Qy 601 AAAGAAAGTTGCCAAAGAGCAGCAAGAGATTTAGAACAAAGAGCTGTATAGCAAAAAA 660

Db 598 AAAGAAAGTTGCCAAAGAGCAGCAAGAGATTTAGAACAAAGAGCTGTATAGCAAAAAA 657

Qy 661 AATTAGAAAGAAAGAAAGAAATGAGATATATTTAGAGAGATTTATATGTCGTTTA 720

Db 658 AATTAGAAAGAAAGAAAGAAATGAGATATATTTAGAGAGATTTATATGTCGTTTA 717

Qy 721 GAAATACAGCTATAGAACTTCCATCAGAAAGAAAGAGTGGATTTATATACCATCAA 780

Db 718 GAAATACAGCTATAGAACTTCCATCAGAAAGAAAGAGTGGATTTATATACCATCAA 777

Qy 781 TCTTCTTACCTCAGCAACAGAGAGATAGTAGAGATTTCCAGAAATATCTATATATA 840

Db 778 TCTTCTTACCTCAGCAACAGAGAGATAGTAGAGATTTCCAGAAATATCTATATATA 837

Qy 841 GAAAAAACAATAGAGATCTATTTACAAATGTTGAGAGAGAGGATATACATATA 900

Db 838 GAAAAAACAATAGAGATCTATTTACAAATGTTGAGAGAGAGGATATACATATA 897

Qy 901 GAGATCTTGAAGAAAGAAAGAGTGTCTCAATATAACAGAACAAAGAAAGAGATTAATCT 960

Db 898 GAGATCTTGAAGAAAGAAAGAGTGTCTCAATATAACAGAACAAAGAAAGATTAATCT 957

Qy 961 GCTGACATACAAATCATATATAGAGACAGTAATATTTCTGATGTTAAATGATTTCAA 1020

Db 958 GCTGACATACAAATCATATATAGAGACAGTAATATTTCTGATGTTAAATGATTTCAA 1017

Qy 1021 ATAAATAGTAGTAGAGATGAAATAGAGTGTGATATAGACATTTATATATGATGAAGAA 1080

Db 1018 ATAAATAGTAGTAGAGATGAAATAGAGTGTGATATAGACATTTATATATGATGAAGAA 1077

Qy 1081 GAAGATAGTAGAGCTTAGAGCAATTTTAAGCTTATGCAATATGCAATTTCCAGAT 1140

Db 1078 GAAGATAGTAGAGCTTAGAGCAATTTTAAGCTTATGCAATATGCAATTTCCAGAT 1137

Qy 1141 GAAGAAACATATAGAAATTTTAAAGAACTAGAGATTTGATAGAGAAAGAAAGAAATTTTA 1200

Db 1138 GAAGAAACATATAGAAATTTTAAAGAACTAGAGATTTGATAGAGAAAGAAAGAAATTTTA 1197

APPLICANT: DRUIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22133-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-32

Query Match 64.1%; Score 950; DB 3; Length 950;
Best Local Similarity 100.0%; Pred. No. 5.8e-155; Mismatches 0; Indels 0; Gaps 0;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

11 AAAGCGATCTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 70
1 AAAGCGATCTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 60
71 TAGAACAGATAGACTTCTTAAGAAAAGTTCAAGAGCGAAAGCGATTAGAACAG 130
61 TAGAACAGATAGACTTCTTAAGAAAAGTTCAAGAGCGAAAGCGATTAGAACAG 120
131 AGAGACTTGTAAAGAAAAGTTGCAAGAACAAAGCGATCTTAAGAACAGAGAGCGTG 190
121 AGAGACTTGTAAAGAAAAGTTGCAAGAACAAAGCGATCTTAAGAACAGAGAGCGTG 180
191 CTAAGAAAAGTTGCAAGAACAAAGCGATCTTAAGAACAGAGAGCGTCTAAAGAAA 250
181 CTAAGAAAAGTTGCAAGAACAAAGCGATCTTAAGAACAGAGAGCGTCTAAAGAAA 240
251 AGTTGCAAGAACAAAGCGATCTTAAGAACAGATAGACTTCTTAAGAAAAGTTACAG 310
241 AGTTGCAAGAACAAAGCGATCTTAAGAACAGATAGACTTCTTAAGAAAAGTTACAG 300
311 AGCAGCAAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAGAA 370
301 AGCAGCAAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAGAA 360

371 GCGATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAG 430
361 GCGATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAG 420
431 AACAGAGAGACTTGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 490
421 AACAGAGAGACTTGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 480
491 GACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTA 550
481 GACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTA 540
551 AAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTTAAGAAAAGT 610
541 AAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTTAAGAAAAGT 600
611 TGCAGAGCGAGCAAGAGATTTAGAACAAAGAGGCTGATACGAAAAAAATTTAGAAA 670
601 TGCAGAGCGAGCAAGAGATTTAGAACAAAGAGGCTGATACGAAAAAAATTTAGAAA 660
671 GAAAAAAGAACATGAGATATATTAAGCAGAGATTTATATGCTTTAGAAAATACAG 730
661 GAAAAAAGAACATGAGATATATTAAGCAGAGATTTATATGCTTTAGAAAATACAG 720
731 CTATAGACTTCCATCAGAAAATGAAACGTGATTTATATACACATCAATCTTCTTAC 790
721 CTATAGACTTCCATCAGAAAATGAAACGTGATTTATATACACATCAATCTTCTTAC 780
791 CTCAGAGCAACAGAGGAATAGTAGAGATTCCAAGGAATATCTATATAGAAAACAA 850
781 CTCAGAGCAACAGAGGAATAGTAGAGATTCCAAGGAATATCTATATAGAAAACAA 840
851 ATAGAGATCTATTACACAAATGTTGAAGAGCAAGAGGATATATTAAGAGACATCTTG 910
841 ATAGAGATCTATTACACAAATGTTGAAGAGCAAGAGGATATATTAAGAGACATCTTG 900
911 AAGAAAAGAAAAGTGGTTCAATTAACACAGAACAAAAAGAAAGTAATCT 960
901 AAGAAAAGAAAAGTGGTTCAATTAACACAGAACAAAAAGAAAGTAATCT 950

RESULT 9
US-08-257-073-15
Sequence 15, Application US/08257073
Patent No. 5766597
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: de Taisne, Charles
TITLE OF INVENTION: MALARIA RECOMBINANT POXYVIRUS VACCINE
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P. C.
STREET: 530 Fifth Avenue, 25th Floor
CITY: New York
STATE: New York
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/257,073
FILING DATE: 09-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/075,783
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/852,305

;; FILING DATE: 18-MAR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/672,183
;; FILING DATE: 20-MAR-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Frommer, William S.
;; REGISTRATION NUMBER: 25,506
;; REFERENCE/DOCKET NUMBER: 454310-2570
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 840-3333
;; TELEFAX: (212) 840-0712
;; TELEX: 425066 CURTMS
;; INFORMATION FOR SEQ ID NO: 15:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1320 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-257-073-15

Query Match 57.6%; Score 853.2; DB 2; Length 1320;
Best Local Similarity 87.9%; Pred. No. 2,7e-138;
Matches 930; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 425 ATTTAGAACAAAGAGACTTGTCTTAAAGAAAGTTGCAAGACAAAGCGATTAGAAC 484
DB 263 ATGTAAAAATGTCTCAACAAATTTCAAAAGTCTTTTAAAGAAATCTTGCTTTCKG 322
QY 485 AAGAGAGCGTCTTAAAGAAAGTTGCAAGACAAAGCGATTAGAACAAAGAGAC 544
DB 323 AQAATATATCTCTTAAAGAAATTAATTAAGAAAGGAATTAATTAAGAACATTA 382
QY 545 GTGCTAAAGAAAGTTGCAAGACAAAGCGATTAGAACAAAGAGCGTCTTAAG 604
DB 383 TAAATGATGATGACGATTAATAAAAAATATATTAAGGCGAAGCGAAAAAGCAAGAG 442
QY 605 AAAAGTTGCAAGAGCGCAAGAGATTAGAACAAAGAGCGTCTTAAGAAAAAAT 664
DB 443 ATCTTGAAGAAAAAGCGCGCGCTTAAAGAAAGCGAAGGCTGATTAAGAAAAAAT 502
QY 665 TAGAAAGAAAAAGAAACATGAGATATATTAGCAGAGATTTATATGCTTTAGAAA 724
DB 503 TAGAAAGAAAAAGAAACATGAGATATATTAGCAGAGATTTATATGCTTTAGAAA 562
QY 725 TACACCTATAGAACTTCCATCAGAAATGAACGTGATATTTATATCAATCAATCTT 784
DB 563 TACCACTATAGAACTTCCATCAGAAATGAACGTGATATTTATATCAATCAATCTT 622
QY 785 CTTTACCTCAGGCAACAGAGGAATATAGAGATTTCAAAGAAATATCTAATATGAAA 844
DB 623 CTTTACCTCAGGCAACAGAGGAATATAGAGATTTCAAAGAAATATCTAATATGAAA 682
QY 845 AAACAATAGAGATCTATTTACAACAAAATGTTGAAGAGCAAGAGATATACATTAAGAC 904
DB 683 AAACAATAGAGATCTATTTACAACAAAATGTTGAAGAGCAAGAGATATACATTAAGAC 742
QY 905 ATCTTGAAGAAAAAGAGATGTTCAATTAACCAAGACAAAGAAAGATTAATCTGCTG 964
DB 743 ATCTTGAAGAAAAAGAGATGTTCAATTAACCAAGACAAAGAAAGATTAATCTGCTG 802
QY 965 AACTACAAATCTATCTTATAGACAGTAAATTTCTGATGTTATATGATTTTCAATTA 1024
DB 803 AACTACAAATCTATCTTATAGACAGTAAATTTCTGATGTTATATGATTTTCAATTA 862
QY 1025 GTAAATGATGAGTGAATTAAGTGTGATATATGACGATTCATTAATATAGAGAGAG 1084
DB 863 GTAAATGATGAGTGAATTAAGTGTGATATATGACGATTCATTAATATAGAGAGAG 922
QY 1085 ATGATGAAGACTTGAAGCAATTTAAGCTTATTTGCAATATGCAATTTCAAGATGAAG 1144
DB 923 ATGATGAAGACTTGAAGCAATTTAAGCTTATTTGCAATATGCAATTTCAAGATGAAG 982
QY 1145 AAAACATGAATTTATTAAGAACTAGAGATTTGATAGAGAAAAATGAATTTAGATG 1204

DB 983 AAAACATGAATTTATTAAGAACTAGAGATTTGATAGAGAAAAATCAAAATTTAGATG 1042
QY 1205 ATTTAGATGAAGAAATGAATAATCATCAGAGAAATTAATCTGAGAGAAAAATTAAG 1264
DB 1043 ATTTAGATGAAGAAATGAATAATCATCAGAGAAATTAATCTGAGAGAAAAATTAAG 1102
QY 1265 GAAAGAAATATGAAGAAAAAGAGATTAATTTTAAACCAATGATTAAGATTTGATG 1324
DB 1103 GAAAGAAATATGAAGAAAAAGAGATTAATTTTAAACCAATGATTAAGATTTGATG 1162
QY 1325 ATGAGCATATTAATAAATATATAAATGATTAAGCGGTTAATTAAGAGAAAAATTTCA 1384
DB 1163 ATGAGCATATTAATAAATATATAAATGATTAAGCGGTTAATTAAGAGAAAAATTTCA 1222
QY 1385 TAAATCATTTCTTCAATATATTGACGAGGCAATGAATTTTCAATCGTGAATGAGT 1444
DB 1223 TAAATCATTTCTTCAATATATTGACGAGGCAATGAATTTTCAATCGTGAATGAGT 1282
QY 1445 TATCTGAAGATATTAATAATATTTTATGAACCTTA 1482
DB 1283 TATCTGAAGATATTAATAATATTTTATGAACCTTA 1320

RESULT 10
US-08-098-327E-34
; Sequence 34; Application US/08098327E
; Patent No. 6270771
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATITIS STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P. O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098,327E
; FILING DATE: 24-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 988 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PUBLICATION INFORMATION:
; PUBLICATION NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; US-08-098-327E-34

ADDRESSEE: Burns, Doane, Swecker & Mathie
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 954 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..954
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-37

Query Match 31.8%; Score 470.8; DB 3; Length 954;
Best Local Similarity 95.7%; Pred. No. 1.6e-72;
Matches 484; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 77 AAGATGACTTGTCTAAAGAAAAGTTCAAGAGCAAGAACGATTTAGAACAAAGAGAC 136
DB 449 AAGAAAACACGCTAAAGAAAAGTTACAGGGCAACAAAGCGATTTCAAGAACAAAGAGAC 508
QY 137 TTGCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAAAGAGACGCTCTAAG 196
DB 509 GTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTCTAAG 568
QY 197 AAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTCTAAGAAAAGTTGC 256
DB 569 AAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTCTAAGAAAAGTTGC 628
QY 257 AAGAACAAAGAACGATTTAGAACAAAGTAGACTTCTCTAAGAAAAGTTACAAAGAGAC 316
DB 629 AAGAACAAAGAACGATTTAGAACAAAGTAGACTTCTCTAAGAAAAGTTGCAAGAAC 688
QY 317 AAGAGCATTTAGAACAAAGAGACGCTCTAAGAAAAGTTGCAAGAACAAAGCGATT 376
DB 689 AAGAGCATTTAGAACAAAGAGACGCTCTAAGAAAAGTTGCAAGAACAAAGCGATT 748
QY 377 TAGAACAAAGAGACGCTCTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAG 436
DB 749 TAGAACAAAGAGACGCTCTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAG 808
QY 437 AGAGACTTGTCTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTG 496
DB 809 AGAGACTTGTCTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTG 868
QY 497 CTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTCTAAGAAA 556

DB 869 CTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTCTAAGAAA 928
QY 557 AGTGCAAGAACAAAGCGATTTA 582
DB 929 GTTGCAAGAACAAAGCGATTTA 954

RESULT 13
US-08-462-625-37
Sequence 37, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathie
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 954 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..954
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-37

Query Match 31.8%; Score 470.8; DB 3; Length 954;
Best Local Similarity 95.7%; Pred. No. 1.6e-72;
Matches 484; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 77 AAGATGACTTGTCTAAGAAAAGTTACAGAGCAAGCGATTTAGAACAAAGAGAC 136
DB 449 AAGAAAACACGCTAAAGAAAAGTTACAGGGCAACAAAGCGATTTCAAGAACAAAGAGAC 508
QY 137 TTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTCTAAG 196

Db 509 GTGCTAAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGACTTGCTAAAG 568
Qy 197 AAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGCGTCTAAAGAAAGTTGC 256
Db 569 AAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGCGTCTAAAGAAAGTTGC 628
Qy 257 AAGAACAAAGAGGATTATTAGAACAGAGAGCGTCTAAAGAAAGTTGCAGAGAGC 316
Db 629 AAGAACAAAGAGGATTATTAGAACAGAGAGCGTCTAAAGAAAGTTGCAGAGAGC 688
Qy 317 AAGCGATTATTAGAACAGAGAGCGTCTAAAGAAAGTTGCAGAGAGCATT 376
Db 689 AAGCGATTATTAGAACAGAGAGCGTCTAAAGAAAGTTGCAGAGAGCATT 748
Qy 377 TAGAACAGAGAGAGCGTCTAAAGAAAGTTGCAGAGAGCATTAGAACAG 436
Db 749 TAGAACAGAGAGAGCGTCTAAAGAAAGTTGCAGAGAGCATTAGAACAG 808
Qy 437 AGAGACTTGCTAAAGAAAGTTGCAGAGAGCATTAGAACAGAGAGAGC 496
Db 809 AGAGACTTGCTAAAGAAAGTTGCAGAGAGCATTAGAACAGAGAGAGC 868
Qy 497 CTAAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGC 556
Db 869 CTAAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGC 928
Qy 557 AGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGAGAGAGAGAG 582
Db 929 GGTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGAGAGAGAGAG 954

RESULT 14
US-09-902-540-1280
; Sequence 1280, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1280
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(1039)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

Query Match 8.8%; Score 129.8; DB 3; Length 1039;
Best Local Similarity 47.2%; Pred. No. 7.4e-14;
Matches 493; Conservative 0; Mismatches 544; Indels 8; Gaps 3;
Qy 306 ACAAGACGACGAACGATTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAGCA 365
Db 2 AAATTAAGCTTC 61
Qy 366 ACAAGACGATTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAGCAACAAAGCA 425
Db 62 AAATTAAGCTTC 121
Qy 426 TTAGAACAGAGAGAGCTTAAAGAAAGTTGCAAGAGCAACAAAGCGATTAGAGCA 485
Db 122 TAAAAAAAAAATTAAGCAATTAAGAAAGTTGCAAGAGCAACAAAGCGATTAGAGCA 181

Qy 486 AGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGC 545
Db 182 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAAGAAAGTTGCAAGAGAGC 241
Qy 546 TGCTAAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGCGTCTAAAGCA 605
Db 242 TATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGCGTCTAAAGCA 301
Qy 606 AAGTTGCAAGAGAGAGGATTATTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAG 665
Db 302 AAAAAAAAAAATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGC 361
Qy 666 AGAAGAAAGAAAGCAATGAGATATTAGAGAGAGATTATTAGAGAGAGATTAGAGAG 725
Db 362 AAAAAAAAAAAGCAATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAG 421
Qy 726 ACCAGCTATGAACTTCATCAGAAATGAAAGCGTCTAAAGAAAGTTGCAAGAGAGC 785
Db 422 AAAAAAAAAAAGCAATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAG 480
Qy 786 TTTACTCAGAGACAGAGAGAGTATGAGAGATTCAGAGAAATCTTAATAGAGAA 845
Db 481 ATTAATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGAGAG 540
Qy 846 AACAATTAAGAGATCTATTCAACAAATGTTGAAGAGAGAGAGATTAAGAGAGCA 905
Db 541 AACAATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGAGAGAG 600
Qy 906 TCTTGAAGAAAGAGATGTTCAATTAAGAACAGAGAAAGAGATTAAGAGAGAGAG 965
Db 601 AAAAAAAAAAAGCAATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAG 660
Qy 966 CATACAAATCATCATTAAGAGAGAGTAAATTTGATGATGATGATGATGATGATG 1025
Db 661 AAAAAAAAAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGAGAGAGAG 714
Qy 1026 TAAGTATGAGAGATTAAGTCTGATGATGATGATGATGATGATGATGATGATGAG 1085
Db 715 AAAAAAAAAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGAGAGAGAG 774
Qy 1086 TGATGAGAGATTAAGAGATTAAGTCTGATGATGATGATGATGATGATGATGATG 1145
Db 775 CAAAAAAAAAATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGAG 834
Qy 1146 AATCATAGAGATTAATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAGAG 1204
Db 835 AATGAGAGAGATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAGAGAGAGAG 894
Qy 1205 ATTATGAGAGATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAGAGAGAGAG 1264
Db 895 AAAAAAAAAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGAGAGAGAG 954
Qy 1265 GAAAGAAATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGAGAG 1324
Db 955 AAAAAAAAAAAGTTGCAAGAACAAAGCGATTATTAGAACAGAGAGAGAGAGAGAG 1014
Qy 1325 ATGAGCATTAATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAGAGAGAG 1349
Db 1015 AAAAAAAAAAATTAAGAAAGTTGCAAGAACAAAGCGATTATTAGAGAGAGAGAG 1039

RESULT 15
US-09-902-540-1357
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 1357
 ; LENGTH: 612
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(612)
 ; OTHER INFORMATION: unsure at all n locations
 US-09-902-540-1357

Query Match 8.2%; Score 120.8; DB 3; Length 612;
 Best Local Similarity 50.3%; Pred. No. 2.5e-12;
 Matches 293; Conservative 0; Mismatches 289; Indels 0; Gaps 0;

QY	808	AATAGTAGAGATTCCAAAGAAATATCTATATAGAAAAACAATAGAAATCTATTACA	867
DB	9	AA	68
QY	868	ACAAATGTTGAAGACGAAGGATATACATAAAGACATCTTGAGAAAGAAAGATGCT	927
DB	69	AA	128
QY	928	TCATATTAACCAACAAAGAAAGATTAATCTGCTGACATCAAAATCATACATTAGAG	987
DB	129	AA	188
QY	988	ACAGTAAATATTTCTGATGATTAATGATTTTCAAAATAGTAAATGATGAGATTAAGT	1047
DB	189	AA	248
QY	1048	GCTGAATATGACGATTCATTATAGATGAAGAAAGATGATGAGACTTAGACGATTT	1107
DB	249	AA	308
QY	1108	AAGCCTATGTGCAATATGCAATTTCCAAAGTGAAGAAACATNGAATTTATTAAGAA	1167
DB	309	AAANATTAATTAATAAAAAAAAAAAAAAAAAATTAATAATAATAATAATAATA	368
QY	1168	CTAGAAGATTTGTAGAGAAAAATGAATAATTTAGATGATTTAGAGAAATAGAAAA	1227
DB	369	ATTAATAAAAAAAAAATTAATTAATAATAATAATAATAATAATAATAATAATA	428
QY	1228	TCATCAGAAAGATTATCTGAGAAAAATTAATAATAATAATAATAATAATAATA	1287
DB	429	AAAAAAAAAAAAAAAAATTAATAATAATAATAATAATAATAATAATAATAATA	488
QY	1288	GATTAATTAATTTTAAACCAATGATTAATAAGTTTGTATGATGAGCATTAATAATA	1347
DB	489	AA	548
QY	1348	AATGATAGAGGTTAATAGAAAGAAAGAAATTCATAAAA	1389
DB	549	AAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	590

Search completed: December 7, 2005, 03:40:57
 Job time : 302 secs

PT vaccination against, treatment of and diagnosis of malaria.
XX
PS Claim 2, Fig 8-10; 81pp; French.

XX This nucleotide sequence is the 3' part of the P.falciparum liver-stage
CC specific antigen (LSA) gene. It codes for a polypeptide sequence which
CC carries a T cell epitope characteristic of a protein produced in
CC hepatocytes infected with P.falciparum. The polypeptide can be used in
CC the preparation of vaccines against malaria. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX

Sequence 1496 BP; 716 A; 169 C; 300 G; 311 T; 0 U; 0 Other;

Query Match 100.0%; Score 1482; DB 2; Length 1496;
Best Local Similarity 100.0%; Pred. No. 8,5e-202;
Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 CAAGAACAAACAAAGCGATCTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCAGAACAA 60
DB 1 CAAGAACAAACAAAGCGATCTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCAGAACAA 60
OY 61 CAAAGCGATTTAGAACAGATGACTTGTCTAAAGAAAAGTTGCAGAGCAGCAAGCGAT 120
DB 61 CAAAGCGATTTAGAACAGATGACTTGTCTAAAGAAAAGTTGCAGAGCAGCAAGCGAT 120
OY 121 TTAGAACAAAGAGACTTGTCTAAAGAAAAGTTGCAGAACAAAGCGATCTAGAACAA 180
DB 121 TTAGAACAAAGAGACTTGTCTAAAGAAAAGTTGCAGAACAAAGCGATCTAGAACAA 180
OY 181 GAGAGACGCTCTAAAGAAAAGTTGCAGAACAAAGCGATTTAGAACAGAGAGACGT 240
DB 181 GAGAGACGCTCTAAAGAAAAGTTGCAGAACAAAGCGATTTAGAACAGAGAGACGT 240
OY 241 GCTAAAGAAAAGTTGCAGAACCAAGCGATTTGAAACAGATGACTTGTCTAAAGAA 300
DB 241 GCTAAAGAAAAGTTGCAGAACCAAGCGATTTGAAACAGATGACTTGTCTAAAGAA 300
OY 301 AAGCTTCAAGAGACGAGAAAGCGATTTAGAACAGAGAGACGCTCTAAAGAAAAGTTGCA 360
DB 301 AAGCTTCAAGAGACGAGAAAGCGATTTAGAACAGAGAGACGCTCTAAAGAAAAGTTGCA 360
OY 361 GAACAAACAAAGCGATTTAGAACAGAGAGACGCTCTAAAGAAAAGTTGCAGAACCA 420
DB 361 GAACAAACAAAGCGATTTAGAACAGAGAGACGCTCTAAAGAAAAGTTGCAGAACCA 420
OY 421 ACGGATTTAGAACAGAGACCTTGTCTAAAGAAAAGTTGCAGAACCAAGCGATTTA 480
DB 421 ACGGATTTAGAACAGAGACCTTGTCTAAAGAAAAGTTGCAGAACCAAGCGATTTA 480
OY 481 GAACAAAGAGACGCTCTAAAGAAAAGTTGCAGAACCAAGCGATTTAGAACAGAG 540
DB 481 GAACAAAGAGACGCTCTAAAGAAAAGTTGCAGAACCAAGCGATTTAGAACAGAG 540
OY 541 AGACGCTCTAAAGAAAAGTTGCAGAACCAAGCGATTTAGAACAGAGACGCTGT 600
DB 541 AGACGCTCTAAAGAAAAGTTGCAGAACCAAGCGATTTAGAACAGAGACGCTGT 600
OY 601 AAAAGAAAAGTTGCAGAGCAGCAAGAGATTGAAACAAAGAGGCTGATCGAAAAA 660
DB 601 AAAAGAAAAGTTGCAGAGCAGCAAGAGATTGAAACAAAGAGGCTGATCGAAAAA 660
OY 661 AATTTGAAAAGAAAAGAGAACTGAGATATTTGACAGAGATTTATATGATCGTTT 720
DB 661 AATTTGAAAAGAAAAGAGAACTGAGATATTTGACAGAGATTTATATGATCGTTT 720
OY 721 GAAATACAGCTATAGAACTTCCATCAGAAAATGAAAGTGATATATATCCACATCA 780
DB 721 GAAATACAGCTATAGAACTTCCATCAGAAAATGAAAGTGATATATATCCACATCA 780
OY 781 TCTTCTTTTACCTGAGACCAACAGAGGAAATGATGAGATTCAGAGAAATATCTATATA 840
DB 781 TCTTCTTTTACCTGAGACCAACAGAGGAAATGATGAGATTCAGAGAAATATCTATATA 840
```

```
OY 841 GAAAAAACAAATAGAAATCTATTACAAACAAATGTTGAGAGACGAAGGATATACATATA 900
DB 841 GAAAAAACAAATAGAAATCTATTACAAACAAATGTTGAGAGACGAAGGATATACATATA 900
OY 901 GACACTCTTTGAGAAAAGAGTGGTTCATTAACCCAGAACCAAAAAGAAATAACT 960
DB 901 GACACTCTTTGAGAAAAGAGTGGTTCATTAACCCAGAACCAAAAAGAAATAACT 960
OY 961 GCTGACATACAAATCATACATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCA 1020
DB 961 GCTGACATACAAATCATACATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCA 1020
OY 1021 ATTAAGTATGAGAGATGAATAATAGTCTGATATGACGATTCATTAATAGTAAAGA 1080
DB 1021 ATTAAGTATGAGAGATGAATAATAGTCTGATATGACGATTCATTAATAGTAAAGA 1080
OY 1081 GAAAGTATGAGAGACTTGAAGCAATTTAAGCTATGTGCAATATGACAAATTTCCAAGAT 1140
DB 1081 GAAAGTATGAGAGACTTGAAGCAATTTAAGCTATGTGCAATATGACAAATTTCCAAGAT 1140
OY 1141 GAAAGAAAACATAGAAATTTATTAAGACCTAGAAAGTTGATPAGAGAAAATGAATTTA 1200
DB 1141 GAAAGAAAACATAGAAATTTATTAAGACCTAGAAAGTTGATPAGAGAAAATGAATTTA 1200
OY 1201 GATGATTTAGATGAAGAAATGAAAAATCATCAGAAATTTATCTGAAGAAAAATTA 1260
DB 1201 GATGATTTAGATGAAGAAATGAAAAATCATCAGAAATTTATCTGAAGAAAAATTA 1260
OY 1261 AAAGAAAAGAAATATGAAAAAACAAGATATAATTTTAAACCAAAAGATATAAGTTTG 1320
DB 1261 AAAGAAAAGAAATATGAAAAAACAAGATATAATTTTAAACCAAAAGATATAAGTTTG 1320
OY 1321 TATGATGACATATTAATAAATATTAATAATGTAAGCAGTTATPAGAAAAAGAAAA 1380
DB 1321 TATGATGACATATTAATAAATATTAATAATGTAAGCAGTTATPAGAAAAAGAAAA 1380
OY 1381 TTCAATTAATCATTTGTTTATATATTTTGAACGAGACAAATGAATTTTACAGATCGTGAT 1440
DB 1381 TTCAATTAATCATTTGTTTATATATTTTGAACGAGACAAATGAATTTTACAGATCGTGAT 1440
OY 1441 GAGTTATCTGAAGATATTAATTAATTTTATGAAACCTRTA 1482
DB 1441 GAGTTATCTGAAGATATTAATTAATTTTATGAAACCTRTA 1482
```

RESULT 2
AAQ28115
ID AAQ28115 standard; DNA; 950 BP.
XX
AC AAQ28115;
XX
DT 25-MAR-2003 (revised)
XX
DT 08-FEB-1993 (first entry)
XX
DE P.falciparum LSA-R-NR coding sequence.
XX
KW Malaria; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;
XX
KW paludism; liver stage-specific antigen; 88.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT repeat_region 3..629 /*tag= a
FT FT /*tag= TANDEM
FT repeat_unit 3..53
FT FT /*tag= b
XX
PN MO9213884-A1.
XX
XX 20-AUG-1992.
XX
XX 05-FEB-1992; 92WO-FR000104.

XX 05-FEB-1991; 91FR-00001286.
XX (INSP) INST PASTEUR.
XX
XX Querimarchand C, Drulhe P;
XX WPI, 1992-299985/36.
XX P-PSDB; AAR26941.
XX
XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX vaccination against, treatment of and diagnosis of malaria.
XX
XX Claim 17, Fig 2, 81pp; French.
XX
XX A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
XX gIII was used to transform E.coli. The expression library was screened
XX with human antiere antigens of all stages of P. falciparum. The
XX library was rescreened with antibodies affinity- purified on a clone
XX which was able to recognise antibodies specific to the hepatic phase.
XX About 40 clones were detected which produced a characteristic LSA
XX epitope. The clone with the largest insert (950 bases) encoded LSA-R-NR
XX containing a 12-repeat region followed by a non-repeat region. Preferred
XX antigenic polypeptides of the invention are derived from the amino acid
XX sequence of LSA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 950 BP; 456 A; 127 C; 206 G; 161 T; 0 U; 0 Other;
Query Match 64.1%; Score 950; DB 2; Length 950;
Best Local Similarity 100.0%; Pred. No. 2,9e-126;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 AAAGGATCTAGAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATT 70
DB 1 AAAGGATCTAGAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATT 60
QY 71 TAGAACAGATGACTTCTTAAGAAAGTTTCAAGAGACCAAGCGATTGAAACAG 130
DB 61 TAGAACAGATGACTTCTTAAGAAAGTTTCAAGAGACCAAGCGATTGAAACAG 120
QY 131 AGAGACTGCTAAAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAGAGACGTTG 190
DB 121 AGAGACTGCTAAAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAGAGACGTTG 180
QY 191 CTAAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAGAGACGTTGCTAAAGAA 250
DB 181 CTAAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAGAGACGTTGCTAAAGAA 240
QY 251 AGTTGCAAGAACAAAGCGATTGAAACAGATGACTTCTTAAGAAAGTTTCAAG 310
DB 241 AGTTGCAAGAACAAAGCGATTGAAACAGATGACTTCTTAAGAAAGTTTCAAG 300
QY 311 AGCAGCAAGGATTGAAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAA 370
DB 301 AGCAGCAAGGATTGAAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAA 360
QY 371 GCGATTGAAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTG 430
DB 361 GCGATTGAAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTG 420
QY 431 AACAGAGAGACTTCTTAAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAG 490
DB 421 AACAGAGAGACTTCTTAAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAG 480
QY 491 GACGCTTAAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAGAGAGCGTCTA 550
DB 481 GACGCTTAAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAGAGAGCGTCTA 540
QY 551 AAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAGAGAGCGTCTAAAGAAAGT 610
DB 541 AAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAGAGAGCGTCTAAAGAAAGT 600
QY 611 TGCAAGAGCAGCAAGAGATTGAAACAAAGAGCGTCTATCGAAAAAAATTTAGAA 670

DB 601 TGCAAGAGCAGCAAGAGATTGAAACAAAGAGCGTCTATCGAAAAAAATTTAGAA 660
QY 671 GAAAAAGAAACATGAGATATATTAGCAGAGATTTATGCTGTTTGAATACAG 730
DB 661 GAAAAAGAAACATGAGATATATTAGCAGAGATTTATGCTGTTTGAATACAG 720
QY 731 CTAATGAACTTCATCGAAGAAATGAACGTGATATATTATACACATCAATCTCTTAC 790
DB 721 CTAATGAACTTCATCGAAGAAATGAACGTGATATATTATACACATCAATCTCTTAC 780
QY 791 CTCAGACAAACAGAGGAAATAGATGATTCCAAAGAAATATCTATATAGAAAAACAA 850
DB 781 CTCAGACAAACAGAGGAAATAGATGATTCCAAAGAAATATCTATATAGAAAAACAA 840
QY 851 ATAGAGATCTATTACAAATGTTGAGAGCAGAAAGGATATACATTAAGACATCTTG 910
DB 841 ATAGAGATCTATTACAAATGTTGAGAGCAGAAAGGATATACATTAAGACATCTTG 900
QY 911 AAGAAAAAGAAAGATGCTTCAATTAACCAAGAAAAAGAAATTAATCT 960
DB 901 AAGAAAAAGAAAGATGCTTCAATTAACCAAGAAAAAGAAATTAATCT 950
RESULT 3
AAQ80916
ID AAQ80916 standard; cDNA; 1320 BP.
XX
XX AAQ80916;
XX AC
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 29-AUG-1995 (first entry)
XX
XX Plasmodium falciparum LSA-1 repeatless gene cDNA.
DE
XX
XX Plasmodium falciparum LSA-1 repeatless gene; recombinant poxvirus;
KW multicomponent multistage malarial vaccines; immunogens;
KW malaria diagnosis; ss.
XX
XX OS Plasmodium falciparum; (PLSARPLS.14L.1).
XX PN W09428930-A1.
XX
XX 22-DEC-1994.
PD
XX
XX 10-JUN-1994; 94MO-US006652.
PF
XX
XX 11-JUN-1993; 93US-00075783.
PR 09-JUN-1994; 94US-00257073.
XX
XX (VIRO-) VIROGENETICS CORP.
PA
XX
XX Paolletti E, De Taisne C, Tine JA;
XX WPI; 1995-036113/05.
XX
XX Recombinant poxvirus contg. Plasmodium DNA in non-essential region -
PT useful in vaccines against malaria and for prodn. of Plasmodium
PT immunogens.
XX
XX Claim 3; Fig 11; 183pp; English.
XX
XX AAQ80916 is the P. falciparum LSA-1 repeatless gene cDNA sequence. New
XX recombinant poxviruses containing either the SERA, ABRA, Pfhp70, AMA-1,
XX Pf25, Pf26, GSP, PfSGP2, LSA-1, LSA-1 repeatless, MSA-1, MSA-1 (N-
XX terminal p83 or C-terminal gp42) genes, or a combination of these in non-
XX essential regions of their genomes are claimed. These poxviruses (pref.
XX with a virulence reducing genomic deletion or disruption) can be used as
XX vaccines against malaria and for the prodn. of Plasmodium immunogens.
XX These viruses provide multicomponent, multistage vaccines due to their
XX expression of sporozite, liver stage, blood stage and sexual stage
XX proteins. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-

OCT-2003 to standardise OS field)

CC Sequence 1320 BP, 619 A, 128 C, 227 G, 346 T, 0 U, 0 Other:

XX SQ Query Match 57.6%; Score 853.2; DB 2; Length 1320;
Best Local Similarity 87.9%; Pred. No. 1.5e-112;
Matches 930; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 425 ATTTGACACAGAGAGACTTGTCTTAAGAAAAGTTGCAGACACACAAAGCGATTGAAAC 484
DB 263 ATGTAAAAAATGTCTCAACAAACAAATTTCCAAAAGCTTTTAAAGAACTTGTGTTTCG 322
QY 485 AAGAGAGAGCTGTCTTAAGAAAAGTTGCAGACACAAAGCGATTGAAACAGAGAGAC 544
DB 323 AGAATATATCTCTTAAGAAAAGTTAAATTAAGAGAGAGAAATTAATGACACATTA 382
QY 545 GTGCTTAAGAAAAGTTGCAGACACAAAGCGATTGAAACAGAGAGAGCTGTCTTAAG 604
DB 383 TAAATGATGATGCGATTAAGAAAAGTTAAATTAAGAGAGAGCGAAAACAGACAGAAAG 442
QY 605 AAAAGTTGCAGAGCGCAAGAGATTGAAACAAAGAGAGCTGATACGAAAAAAAT 664
DB 443 ATCTTGAGAGAAAAGCGCGCATCTMAAGAAACGAGAGAGCTGATACGAAAAAAAT 502
QY 665 TAGAAAAGAAAAGAAACATGAGATATATTAGCAGAGAGATTATGTGCTTTAGAA 724
DB 503 TAGAAAAGAAAAGAAACATGAGATGTTAGCAGAGAGATTATGTGCTTTAGAA 562
QY 725 TACAGACTATAGAACTTCATCAGAAAATGAACGTGATATTATATACCATCAATCTT 784
DB 563 TACAGACTATAGAACTTCATCAGAAAATGAACGTGATATTATATACCATCAATCTT 622
QY 785 CTTTACTCTCAGACACAAGAGGAAATAGTAGAGATTCCAAAGAAAATCTTATATAGAA 844
DB 623 CTTTACTCTCAGACACAAGAGGAAATAGTAGAGATTCCAAAGAAAATCTTATATAGAA 682
QY 845 AAAACAATAGAGATCTATTACAAACAATGTTGAAGAGAGAGATATACATTAAGAGAC 904
DB 683 AAAACAATAGAGATCTATTACAAACAATGTTGAAGAGAGAGATATACATTAAGAGAC 742
QY 905 ATCTTGAGAAAAGAAAAGATGTTCAATTAACACAGAACAAAAGAAATTAATCTGCTG 964
DB 743 ATCTTGAGAAAAGAAAAGATGTTCAATTAACACAGAACAAAAGAAATTAATCTGCTG 802
QY 965 ACATTAACAATCTATCAATTAAGAGACAGTAATATTTCTGATGTTATGATTTCAAATA 1024
DB 803 ACATTAACAATCTATCAATTAAGAGACAGTAATATTTCTGATGTTATGATTTCAAATA 862
QY 1025 GTAAGATAGAGATGAATTAAGTGCAGATATAGACGATTCATTAATAGATGAAGAGAG 1084
DB 863 GTAAGATAGAGATGAATTAAGTGCAGATATAGACGATTCATTAATAGATGAAGAGAG 922
QY 1085 ATGATGAGAGCTTAGACGAAATTAAGCTATTGTGCAATATAGCAATTTCCAAAGTAG 1144
DB 923 ATGATGAGAGCTTAGACGAAATTAAGCTATTGTGCAATATAGCAATTTCCAAAGTAG 982
QY 1145 AAAACCTAGGAATTTTAAAGAACTAGAGATTTGATAGAGAAAATGAAAAATTTGAG 1204
DB 983 AAAACCTAGGAATTTTAAAGAACTAGAGATTTGATAGAGAAAATGAAAAATTTGAG 1042
QY 1205 ATTTAGATGAAGATAGAGAAAATCATCAGAGAAATTAATCTGAAAGAAAATTAAGAA 1264
DB 1043 ATTTAGATGAAGATAGAGAAAATCATCAGAGAAATTAATCTGAAAGAAAATTAAGAA 1102
QY 1265 GAAAGAAATATGAAAAAACAAGAGATTAATTTTAAACCAATGATAAAAAGTTGTATG 1324
DB 1103 GAAAGAAATATGAAAAAACAAGAGATTAATTTTAAACCAATGATAAAAAGTTGTATG 1162
QY 1325 ATGAGCATATTTAAAAAATATAAAAATGATAGCAGGTTAATTAAGAAAAAGAAAAATTC 1384
DB 1163 ATGAGCATATTTAAAAAATATAAAAATGATAGCAGGTTAATTAAGAAAAAGAAAAATTC 1222
QY 1385 TAAATCATTTGTTCAATATATTGACGAGACAAATGAATTTTACAGATGTGATGAGT 1444

DB 1223 TAAATCATTTGTTCAATATATTGACGAGACAAATGAATTTTACAGATCGTGATGAGT 1282

QY 1445 TATCTGAAGATATTAATTAATTTTATGAAACATTA 1482

DB 1283 TATCTGAAGATATTAATTAATTTTATGAAACATTA 1320

RESULT 4
AA005140
ID AA005140 standard; DNA, 1566 BP.

XX AC AA005140;
XX AC 25-MAR-2003 (revised)
XX DT 05-NOV-1990 (first entry)
XX DE Sequence encoding N-terminal of peptide antigen to malarial sporozite.
XX KM Malaria; sporozite; vaccine; exoerythrocytic parasites; tetanus toxoid.
XX OS Plasmodium falciparum.
XX FT Key Location/Qualifiers
XX FT CDS 169..1554
XX FT /*tag= a
XX PN WO9006130-A.
XX PD 14-JUN-1990.
XX PF 30-NOV-1988; 88US-00278234.
XX PR 30-NOV-1988; 88US-00278234.
XX PR 12-APR-1989; 89US-00357204.
XX PA (BIOM-) BIOMEDICAL RES. INST.
XX PI Hollingdal MR;
XX DR WPI; 1990-209624/27.
XX DR P-PSDB; AAR05766.
XX PT Novel malarial sporozoite antigenic protein - useful as vaccine against
XX PT sporozoite(s) and exo-erythrocytic parasites.
XX PS Disclosure; Page 7; -pp; English.
XX CC Antigen, preferably linked at the C-terminal to a carrier such as tetanus
XX CC toxoid, may be used as a vaccine against the malarial sporozoite. (Updated
XX CC on 25-MAR-2003 to correct PI field.)

SQ Sequence 1566 BP, 745 A, 210 C, 330 G, 281 T, 0 U, 0 Other:

Query Match 39.6%; Score 586.4; DB 2; Length 1566;
Best Local Similarity 91.7%; Pred. No. 1.1e-74;
Matches 620; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 CAAGAACAAACAAGCGATCTAGACACAGAGAGAGCTGCTTAAGAAAAGTTGCAAGACAA 60
DB 745 CAAGAACAAACAAGCGATTTAGAACAGAGAGAGCTGCTTAAGAAAAGTTGCAAGACAA 804
QY 61 CAAGCGATTTGAAACAAGATGACTTGTAAAGAAAAGTTCAAGAGAGCAAGCGAT 120
DB 805 CAAGCGATTTGAAACAAGAGAGAGCGTGTAAAGAAAAGTTGCAAGACAAAGCGAT 864
QY 121 TTAGAACAAAGAGAGAGCTTGCTAAAGAAAAGTTGCAAGACAAAGCGATTAGAACAA 180
DB 865 TTAGAACAAAGAGAGAGCTTGCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAA 924
QY 181 GAGAGAGCTGTAAAGAAAAGTTGCAAGACAAACAAGCGATTTAGACACAGAGAGAGCT 240
DB 925 GATAGAGCTGTAAAGAAAAGTTGCAAGACAAACAAGCGATTTAGACACAGATGAGCTT 984

QY 241 GCTAAGAAAAGTTGCAAGAACCAAGCGATTTAGAACAGATAGACTTGTCTAAGAA 300
DB 985 GCTAAGAAAAGTTGCAAGAACCAAGCGATTTAGAACAGATAGACTTGTCTAAGAA 1044
QY 301 AAGTTACAGAGCAGCAAGCGATTTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAA 360
DB 1045 ACGTTGCAAGAGCCAGCAAGCGATTTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAA 1104
QY 361 GAAACAAAGCGATTTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAAAGAACAA 420
DB 1105 GAAACAAAGCGATTTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAAAGAACAA 1164
QY 421 AGCGATTTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAAAGAACAGGATTTA 480
DB 1165 AGAGATTTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAAAGAACAGGATCTA 1224
QY 481 GAAACAGAGACGTGCTAAAGAAAAGTTGCAAAGAACAGGATTTAGAACAGAG 540
DB 1225 GAAACAGAGACGTGCTAAAGAAAAGTTGCAAAGAACAGGATTTAGAACAGAG 1284
QY 541 AGAGCGCTAAAGAAAAGTTGCAAAGAACAGGATTTAGAACAGAGAGACGTGCT 600
DB 1285 AGACTGCTAAAGAAAAGTTGCAAAGAACAGGATTTAGAACAGAGAGACGTGCT 1344
QY 601 AAAGAAAAGTTGCAAGAGACGCAAGAGATTTAGAACAGAGAGAGCTGATTCGAAAA 660
DB 1345 AAAGAAAAGTTGCAAGAGACGCAAGAGATTTAGAACAGAGAGAGCTGATTCGAAAA 1404
QY 661 AATTAGAAAAGAAA 676
DB 1405 TTGCAAGAACAAACAA 1420

RESULT 5

ADO21941
ID ADO21941 standard; DNA; 1371 BP.

XX ADO21941;

AC 12-AUG-2004 (first entry)

XX DE LSA-NRC(H) construct DNA derived from Malaria parasite LSA-1.

XX LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;

XX K1 parasitophorous vacuole; antimalarial; vaccine; immunotherapy;

XX K2 malaria parasite P. falciparum; LSA-NRC(H) construct; ds; gene.

OS Plasmidium falciparum.

XX Synthetic.

XX Key

PH Location/Qualifiers

FT 1..1371

FT /product= "LSA-NRC(H) construct protein derived from

FT Malaria parasite LSA-1"

XX MO200404167-A2.

XX 27-MAY-2004.

XX 12-NOV-2003; 2003WO-US036011.

XX 12-NOV-2002; 2002US-0425719P.

XX (REED-) REED ARMY INST RES WALTER.

XX Laner DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;

XX Barbosa A;

XX WPI; 2004-420309/39.

XX P-PSDB; ADO21942.

PT Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
XX Claim 16; SEQ ID NO 25; 90pp; English.
PS
XX The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct DNA of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to a C-terminal His6 tag.
XX
SQ Sequence 1371 BP; 539 A; 358 C; 283 G; 191 T; 0 U; 0 Other;
Query Match 34.0%; Score 503.4; DB 12; Length 1371;
Best Local Similarity 63.1%; Pred. No. 6.9e-63;
Matches 791; Conservative 0; Mismatches 461; Indels 1; Gaps 1;
QY 227 AACAGAGAGACGTGCTAAGAAAAGTTGCAAGAACCAAGCGATTTAGAACAGAT 286
DB 79 AACAGAGAAAAGCATGAGAGAAAACATGCTGAGCCACAACTCTACGAGAAAGACTAAA 138
QY 287 GACTTGCTAAGAAAAGTTGCAAGAGAGCAAGACGATTTAGAACAGAGAGACGTGCT 346
DB 139 AACAGAGAAAACAAATCTTTGACAAAGAGCAAGAGCTGACGATGACGAACTTTAAA 198
QY 347 AACAGAAAAGTTGCAAGAACCAAGCGATTTAGAACAGAGAGACGTGCTAAGAAAAGTTGCAAGAA -AAG 405
DB 199 AACGATCCAGACCAACTTTAAATCCCTCGGCAACTCGGCTTCCGAGAACATC 258
QY 406 TTGCAAGAACCAAGCGATTTAGAACAGAGAGACGTGCTAAGAAAAGTTGCAAGAA 465
DB 259 TTTCTCAAGAAAACAACTGAAACAGAGAGCAAACTGATTAACATATCATCAAGAC 318
QY 466 CAACAAAGCGATTTAGAACAGAGAGACGTGCTAAGAAAAGTTGCAAGAACCAAGAC 525
DB 319 GAGCATGACAAAATAATCATTAAGGCGAGATGAATAATGCGCAAGACCTCGAA 378
QY 526 GATTAGAACAGAGAGACGTGCTAAGAAAAGTTGCAAGAACCAAGACGATTTGAA 585
DB 379 GAAAAGACGTGTAACGACGACGACCTGGAACAGAGAGCGCTCGCTAAGAAAAGCTC 438
QY 586 CAAGAGAGCGTGTAAAGAAAAGTTGCAAGAGACGCAAGAGATTTAGAACAAAGAGAG 645
DB 439 CAGAGAGCGCTCGCTAAGAAAAGCTCAGAGAGCAAGCGCACTGGAACAGCGAAG 498
QY 646 GCTGATACGAAAAAATTTAGAAAAGAAAAGCAATGAGATATATAGCAGAGAT 705
DB 499 GCTGACAGGAAAAAATCTGGAAGCGCAAAAGGAAACCGCGGCGCTGAGAGAC 558
QY 706 TTATATGCTGTTTGAAGATACAGCTATAGAACTTCAATCAAGAAAAGTGAAGTAT 765
DB 559 CTGTACGCGCGCTGGAATATCCAGCTATGAACTCCCATCCGAAAAGCAAGCGGCTAC 618
QY 766 TATATACCAATCAATCTTTTACCTCAGAGACAAAGAGGAAATATAGATTTCCAG 825
DB 619 TACATCCACACAGAGACGCTGCAAGATATATGCGGGAACCTCCCGGACAGTAG 678
QY 826 GAAATATCTATATAGAAAAAACAATAGAGATCTATTACAAACAAATGTTGAAGAGCA 885
DB 679 GAATTCAGCATCATGAAAAAACAACCGGAAAGCATTTCCACCAAGCGTGAAGCGCG 738
QY 886 AGGATATACATTAAGAGACATCTTGAAGAAAAGAAAAGTGTCAATTAATTAACCAAGCA 945
DB 739 CGGACATCCCAAGAGCGCACTGGAAGAAAAGAAAAGCGCTCCACCAACCAAGCAAG 798
QY 946 AAAGAGATTAATGCTGACATCAAAATCATATTAGAGACAGTAAATATTCTTGAT 1005

```
Db      799  AAGAAGACAAAGCGCTGATATCCAGAACCAACCCCTGGAGACCGTGAAATTAAGCGAC 858
Oy      1006 GTTAATGATTTTCAAAATAGTAAGTAGGATGAAATTAAGCTGAATATGACGATTCA 1065
Db      859  GTGAACGACTTCCAGATCAGCAAGTAGCGAGACGAATCTCCGCTGAATACGATGACTCC 918
Oy      1066 TTAATAGAGAAAGAAATGATGATGAGACTAGACGAATTTAAGCTTTAGCTTTGCGAATAT 1125
Db      919  CTGATCGACGAAAGAAAGACGACGAGAGATCTGAGTAATTTCAAAACCAATTTGTCAGTAC 978
Oy      1126 GACAAATTTCCAGATGAGAGAAACATAGGAATTTTAAAGAACTAGAGATTTGATAGAG 1185
Db      979  GATTAATTTCAAGAGACGAAAGAAATATCGGCATTTTACAAAGAACTGAGAGACCTCATCGAG 1038
Oy      1186 AAAAAATGAAATTTTATGATGATTTAGTAGAGAAATGAAATTTATCATCAGAGAAATTTATCT 1245
Db      1039 AAAAAAGAAACCTGACGACCTGCGAGAGAGATCGAAAAATTCCTCCAGAAACCTGAGC 1098
Oy      1246 GAAAGAAAAATTAAGAAAGAAAGAAATTAAGAAAAACAAAGGATTAATTTTAAACCA 1305
Db      1099 GAAAGAAAAATTAAGAAAGAAAGAAATTAAGAAAAACAAAGGACAACTTCAACCA 1158
Oy      1306 AATGATTAAGATTTTATGATGATGACATATTAATAAATATTAATAATGATAGCAGTTAAT 1365
Db      1159 AAGCAAAATCCCTCTACGACGACGACATTTAAAAATTAACAAAACGACAGCAAGTGAC 1218
Oy      1366 AAGGAAAGGAAAAATTCATTAATATGTTTCAATATTTTGAACGAGACATGAAT 1425
Db      1219 AAGGAAAGGAAAAATTTATCAAAATCCCTTCCTCAATCTTCGATGCGCATTAACGAAT 1278
Oy      1426 TTAAGATGCTGATGATGATGATCTGAAGATATTAACCTAAATTTTAAACCT 1478
Db      1279 CTGCAATATTTAGACGAACTGAGCGAAGACATCACTTAATTAATCTTCAAGAGCT 1331

RESULT 6
AD021919 standard; DNA: 1374 BP.
XX
AC      ADO21919;
XX
DT      12-AUG-2004 (first entry)
XX
DE      LSA-NRC(H)Mut construct DNA derived from Malaria parasite LSA-1.
XX
KW      LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
KW      parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
KW      malaria parasite P. falciparum; LSA-NRC(H)Mut construct; ds; gene;
KW      mutant.
XX
OS      Plasmodium falciparum.
OS      Synthetic.
XX
FH      Key      Location/Qualifiers
FH      CDS      1..1374
FH      FT      /tag= a
FH      FT      /product= "LSA-NRC(H)Mut construct protein derived from
FT      Malaria parasite LSA-1"
XX
XX      MO200404167-A2.
XX
XX      27-MAY-2004.
XX
XX      12-NOV-2003; 2003WO-US036011.
XX
XX      12-NOV-2002; 2002US-0425719P.
XX
XX      (REED-) REED ARMY INST RES WALTER.
XX
XX      Lanat DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W,
XX      Barbosa A,
XX      MPI, 2004-420309/39.
```

```
DR      P-PSDB; AD021920.
XX
PT      Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT      agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT      (LSA-1) epitope.
XX
XX      Claim 17; SEQ ID NO 3; 90pp; English.
XX
CC      The invention relates to a novel recombinant LSA-NRC polypeptide
CC      comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC      protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC      falciparum. The polypeptide of the invention demonstrates antimalarial
CC      activities and may be useful as a vaccine during immunotherapy and as a
CC      diagnostic agent, preferably for diagnosing malaria. The current sequence
CC      is that of the LSA-NRC(H)Mut construct DNA of the invention. The
CC      construct is harmonised for expression in Escherichia coli and comprises
CC      the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC      parasite P. falciparum LSA-1 in addition to an amino acid insertion
CC      within the T5 P. falciparum epitope and a C-terminal His6 tag.
XX
SQ      Sequence 1374 BP; 539 A; 358 C; 285 G; 192 T; 0 U; 0 Other;
XX
Query Match      33.3%; Score 493.6; DB 12; Length 1374;
Best Local Similarity 63.1%; Pred. No. 1.7e-61;
Matches 793; Conservative 0; Mismatches 459; Indels 4; Gaps 2;

227  AACCAAGAGACGCTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAATA 286
Oy      227  AACCAAGAGACGCTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAATA 286
Db      79  AACCGGAAAAGCATGAGAAAGAACATGCTGACCCCAACTCTACGAGAAAGCTAAA 138
Oy      287  GACTGCTAAAGAAAAGTTGCAAGAGCAAGCGATTTAGAACAAAGAGAGCGTGCTA 346
Db      139  AACCAAGAAAACAAATAATCTTTGACAGAGACAAAGAGCTGACATGACGAAAGCTTAA 198
Oy      347  AAGAAAAGTTGCAAGACAAAGCGATTTAGAACAAAGAGACGCTGCTAAAGAA-ANG 405
Db      199  AACGTATCCAGACCAACTTTAAATCCCTCGCCCAACTCGCGGTTTCCGCAACATC 258
Oy      406  TTGCAAGAACAAAGCGATTTAGAACAGAGACCTTCTTAAGAAAAGTTGCAAGAA 465
Db      259  TTTCTCAAGAAAACAACTGAAACAAAGAGCAACTGATTGAACATATCTCAACGAC 318
Oy      466  CAACAAGCGATTTGAAACAAGAGACGCTGTAAGAAAAGTTGCAAGAACAAAGC 525
Db      319  GACGATGACAAAATAATTAATTAAGCCAGATGAAATCCGACAGAAACACTCGAA 378
Oy      526  GATTTAGAACAAAGAGACGCTGTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAA 585
Db      379  GAAAAAGCTGCTGAACAGACGCTGGAACGAGAGCGCTCGCTTAAGAAAAGCTC 438
Oy      586  CAAGAGACGCTGCTAAAGAAAAGTTGCAAGAGCGCAAGAGATTTGAACAAAGAAAG 645
Db      439  CAGGAGCGCTGCTTAAGAAAAGCTCAAGAGCAACGCGGACCTGGAACAGCGCAAG 498
Oy      646  GCTGATACGAAAAAATTTAGAAAAGAAAAGAACTGAGATATATTACAGAGAT 705
Db      499  GCTGACAGAAAAAATCTGGAACGCAAAAAGAAACGCGGACGTTCTGCTAGAGAC 558
Oy      706  TTAATGCTGTTTGAATAATACAGCTATAGAACTTCATCAGAAAATGAACGCTGATAT 765
Db      559  CTGTAAGGCGCGCTGGAATCCAGCTATCGAAATCCCATCCGAAACGAAAGCGGCTAC 618
Oy      766  TATATACCAATCAATCTTTTACCTCAGACCAACAGAGGAATATGATGATTCAGAG 825
Db      619  TACATCCCAACACAGAGACGCTGCAAGATTAATCGGGGAATCCCGCAGACGTAAG 678
Oy      826  GAAATATCTATATAGAAAAAACAATAGAGATCTATTACAACAATTTGAGAGACGA 885
Db      679  GAAATCAGCATATGAAAAAACAACGCGGAAAGCATTTACCAACAGCTGAAAGCGCC 738
Oy      886  AAGGATATATCAATAAAGACATCTTGAAGAAAAGAAAGATGTTCAATAAACAGAAACA 945
Db      739  CGGCAATCCAAAGAGCGCACTCGAAGAAAAGAAAGACGGCTCATCAAAACAGAACAG 798
```

Qy	946	AAAGAAGTAATCTCGACATTCAGAAATCATTCATTAGAGACAGTAAATTTTCAT	1005
Db	799	AAAGAAGCAAAAGCGCTGATATTCAGAACCAACCTCGAGACCTGTAACATTTAGCGAC	858
Qy	1006	GTTAATGATTTTCAAATTAAGTAAGTATGAGATGAGATGAATTAAGTGTCTGATATGACGATTTCA	1065
Db	859	GTGAACGACCTTCAGATCAGCAAGTACGAGGACGAAATCTCGGTGAATTAAGATGACCTCC	918
Qy	1066	TTAATAGATGAAGAAGAAAGATGATGAGACTTAGACGATTTTAAGCCTATTGTGCAATAT	1125
Db	919	CTGATCGACGAAGAAGAGACGACGATCTGATGATTTCAACCAATTTGTCAGTAC	978
Qy	1126	GACAAATTTCCAAAGTGAAGAAACATAGGAATTTATAAGAAGTAAAGATTTGATAGAG	1189
Db	979	GATAACTTTCCAGACGAAAGAAATATTCGGCATTTTCAAGAAGCTCGAAGACCTCATTCGAG	1038
Qy	1186	AAAAATGAAATTTTAGATGATTTAGATGAAGAAATAGAAAAATCATCAGAAATTAATCT	1245
Db	1039	AAAAACGAAACCTCGACGACCTGAGCGAAGGCAATCGAAAAATCCTCGAAGAACTGAGC	1098
Qy	1246	GAGCAAAAAATTAAAAAAGAAAGAAATATGAAAAACAAAGAGATTAATTTTAAACCA	1305
Db	1099	GAAGAAAAATCAAAAAAGCGAAGAAATACAAAAACCAAGGACACACCTTCAAACCA	1158
Qy	1306	AATGATAAAGATTGTATGATGAGCATTTAAAAAATTAATAAAATGATPAAGCAAGTTAAT	1365
Db	1159	AACGACAAATCCCTTACGACGAGCAGCATTAATAAAATCAAAAAAGACAGCAGAGTGAC	1218
Qy	1366	AAGCAAAAGAAAAATTCATTAATCATTTGTTTCATATATTTTGAACGAGACAAAGAAAT	1429
Db	1219	AAGGAAAAAGAAAAATTTATCAAAATCCCTCTTCGACATCTTCGATGGCGATPAAGAAAT	1278
Qy	1426	TTACAGATCGTGATGA--GTTATCTGAAGATATTAATTAATTTTATGAACT	1478
Db	1279	CTGCAAAATGTATGACGAAAGCGTTGAGCGAAGACATCAATAATCTTCATGAAGCT	1334
RESULT 7			
AAQ28117			
ID	AAQ28117	standard; DNA; 988 BP.	
XX	AAQ28117;		
AC			
XX			
DT	25-MAR-2003	(revised)	
DT	08-FEB-1993	(first entry)	
XX			
XX	P.falciparum LSA gene 5' region.		
DE			
XX			
KW	Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;		
KW	paludism; liver stage-specific antigen; ss.		
XX			
OS	Plasmodium falciparum.		
XX			
XX			
PH	Key	Location/Qualifiers	
FT	5'UTR	1..32	
FT		/*tag= a	
FT	misc_feature	33..988	
FT		/*tag= d	
FT		/note= "5' part of LSA gene"	
FT	repeat_region	492..988	
FT		/*tag= b	
FT	repeat_unit	492..542	
FT		/*tag= c	
XX			
PN	W09213884-A1.		
XX			
PD	20-AUG-1992.		
XX			
BP	05-FEB-1992;	92MO-FR000104.	
XX			
PR	05-FEB-1991;	91FR-00001286.	
XX			

Query Match	31.9%; Score 472.8; DB 2; Length 988;
Best Local Similarity	95.7%; Pred. No. 1.5e-58;
Matches	486; Conservative 0; Mismatches 22; Indels 0; Gaps 0
QY	77 AAGATGACCTTGTCTTAAGAAAAAGTTACAAGAGCGCAAAAGCGATTTAGAACAGAGAGAC 136
Db	481 AAGAAAAAGCGCTTAAGAAAAAGTTACAGGGGCAACAAAGCGATTCAAGAACAGAGAGAC 540
QY	137 TTGCTAAAGAAAAGTTGCAAGAACAAACGCGATCTGAACCAAGAGAGAGCGTCTTAAG 196
Db	541 GTGCTAAAGAAAAGTTGCAAGAACCAACGCGATTGAACCAAGAGAGAGCTTCTTAAG 600
QY	197 AAAAGTSCAAGAACCAACAAAGCGATTAGAACAAAGAGAGAGCGTCTAAAGAAAAGTTGC 256
Db	601 AAAAGTSCAAGAACCAACAAAGCGATTAGAACCAAGAGAGAGCGTCTAAAGAAAAGTTGC 660
QY	257 AAGAACCAAAAGCGATTAGAACCAAGATGACCTTGTAAAGAAAAGTTACAAGAGACGC 316
Db	661 AAGAACCAAAAGCGATTAGAACCAAGAGAGAGCTTGTAAAGAAAAGTTGCAAGAACAC 720
QY	317 AAAAGGATTTAGAACCAAGAGAGCGTCTAAAGAAAAGTTCCAGAACCAAGAGCGATT 376
Db	721 AAAAGGATTTAGAACCAAGAGAGCGTCTAAAGAAAAGTTCCAGAACCAAGAGCGATT 780
QY	377 TAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACCAAGCGATTTAGAACCAAG 436
Db	781 TAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACCAAGCGATTTAGAACCAAG 840
QY	437 AGAGACTTGCTTAAGAAAAGTTGCAAGAACCAAGCGATTTAGAACCAAGAGAGCGTG 496
Db	841 AGAGACTTGCTTAAGAAAAGTTGCAAGAGCGCAAGAGCGATTTAGAACCAAGTAGACTTG 900
QY	497 CTAAAGAAAAGTTGCAAGAACCAAGCGATTTAGAACCAAGAGAGAGCGTCTAAAGAAA 556
Db	901 CTAAAGAAAAGTTGCAAGAACCAAGCGATTTAGAACCAAGAGAGAGCGTCTAAAGAAA 960
QY	557 AGTTCAGAGAACCAAGCGATTAGAACCAAGCGATTAGAACCAAGCGATTAGAACCAAG 584
Db	961 GGTTCAGAGAACCAAGCGATTAGAACCAAGCGATTAGAACCAAGCGATTAGAACCAAG 988

RESULT 8
ADP85917

ID ADP85917 standard; DNA; 1300 BP.

XX ADP85917;

XX 26-AUG-2004 (first entry)

DE Synthetic construct #1.

XX Nanopore data analysis ; polymer ; polynucleotide;

QY 1351 GATAAGCAGTTTAAAGAAAAGAAAATTCTATAAA 1389
| | | | | | | | | | | | | | | | | | | | | |
DB 1261 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1299
| | | | | | | | | | | | | | | | | | | | | |
RESULT 9
AA05868 standard; DNA; 3399 BP.
XX
AC AA05868;
XX
DT 27-AUG-2003 (revised)
DT 14-AUG-1996 (first entry)
XX
DE Chicken leucocytozoan DNA encoding immunogenic protein for vaccines.
XX
KM Chicken leucocytozoan; immunogen; recombinant vaccine; protection;
KM Immunisation; vaccination; ss.
XX
OS Leucocytozoan sp.
XX
FH Key Location/Qualifiers
FH CDS 1..3399
FT /*tag= a
FT mlec_feature 1150..3218
FT /note= b
FT /note= "fragment referred to in the claims, for use as
FT insert in a recombinant vaccine against chicken
FT leucocytozoan disease"
XX
PN JP07284392-A.
XX
XX 31-OCT-1995.
XX
XX 19-APR-1994; 94JP-00080643.
XX
XX 19-APR-1994; 94JP-00080643.
XX
XX 19-APR-1994; 94JP-00080643.
XX
XX (DOBU-) DOBUTSUYO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.
XX
XX PA (KITA) KITASATO KENKYUSHO SH.
XX
XX WPI; 1996-006311/01.
XX
XX DR P-PSDB; AAR97866.
XX
XX Chicken leucocytozoan immunogenic protein - used in a recombinant vaccine
XX PT against chicken leucocytozoan disease.
XX
XX PS Claim 6; Page 6-9; 35pp; Japanese.
XX
XX CC AA05868 encodes a chicken leucocytozoan immunogenic protein, this DNA or
XX CC a fragment of it can be used in a recombinant vaccine to immunise against
XX CC chicken leucocytozoan disease. The DNA is used in a vector and
XX CC operatively linked to an expression regulatory sequence as in standard
XX CC practice. (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 3399 BP; 1577 A; 508 C; 798 G; 516 T; 0 U; 0 Other;
SO
Query Match 15.2%; Score 225.8; DB 2; Length 3399;
Best Local Similarity 48.7%; Pred. No. 1.6e-23;
Matches 713; Conservative 0; Mismatches 737; Indels 15; Gaps 3;
QY 2 AAGAACAAACAAGCGATTAGAACAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAC 61
| | | | | | | | | | | | | | | | | | | | | |
DB 1685 AAGTAACATGACGAAAGAAATAGAAAAGAGACATGAAGAAATATCATGAAAGAA 1744
| | | | | | | | | | | | | | | | | | | | | |
QY 62 AAAGGATTTAGAACAAATAGACTTGTCTAAAGAAAGTTTCAAGAGACAGAAAGCGATT 121
| | | | | | | | | | | | | | | | | | | | | |
DB 1745 AAGAAAGAGTAACACATAGAAATAGAAAAGAAAGACATGAAAGTAATACATGAG 1804
| | | | | | | | | | | | | | | | | | | | | |
QY 122 TAGAACAGAGAGACTTCTAAAGAAAAGTTGCAAGAACAAAGCAGCTAGAACAG 181
| | | | | | | | | | | | | | | | | | | | | |
DB 1805 AAGAAAAGAAAGAGTAACACATGAG-----AAGAAAAGAAAGAAAGCAGTGAAG 1858
| | | | | | | | | | | | | | | | | | | | | |

QY 182 AGACAGCTGCTAAAGAAAAGTTGCAAGAACAAACGAATTTAGAACAGAGACGTG 241
| | | | | | | | | | | | | | | | | | | | | |
DB 1859 TAATACATGAAAGAAAAGAAAAGAAAGTAACACATGAAAGAAATAGAAAAGACATG 1918
| | | | | | | | | | | | | | | | | | | | | |
QY 242 CTAAGAAAAGTTGCAAGAACAAACGAATTTAGAACAGATGACTTGTCTAAAGAAA 301
| | | | | | | | | | | | | | | | | | | | | |
DB 1919 AAGAAATTAATCATGAAAGAAAAGAAAGAAAGTAACACATGAAAGAAATAGAAAAG 1978
| | | | | | | | | | | | | | | | | | | | | |
QY 302 AGTTCAAGAGCAGCAAAAGCGATTTAGAACAGAGACGTGCTTAAAGAAAAGTTGCAAG 361
| | | | | | | | | | | | | | | | | | | | | |
DB 1979 AGCATGAAGAAAGTAATCATGAGGAAAAAGAAAGTAACACATGAAAGAAATAGAAA 2038
| | | | | | | | | | | | | | | | | | | | | |
QY 362 AACAAACAAGCGATTTAGAACAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAAACAA 421
| | | | | | | | | | | | | | | | | | | | | |
DB 2039 AAGAA-----GACATGAAAGAACTAATCATGAAAGAAAAGAAAGAAAGTAACATG 2092
| | | | | | | | | | | | | | | | | | | | | |
QY 422 GCGATTTAGAACAGAGACCTGCTTAAAGAAAAGTTGCAAGAACAAACGAATTTAG 481
| | | | | | | | | | | | | | | | | | | | | |
DB 2093 AAGAAATTAAGAAAAGAAAGACATGAAAGTAATCATGAAAGAAAAGAAAGTAAGTA 2152
| | | | | | | | | | | | | | | | | | | | | |
QY 482 AACAGAGAGACGTCTTAAAGAAAAGTTGCAAGAACAAACGAATTTAGAACAGAGA 541
| | | | | | | | | | | | | | | | | | | | | |
DB 2153 CACATGAAAGAAAAGAAAAGAAAGACATGAAAGTAATCATGAAAGAAAAGAAAG 2212
| | | | | | | | | | | | | | | | | | | | | |
QY 542 GACGTGCTTAAAGAAAAGTTGCAAGAACAAACGAATTTAGAACAGAGACGTGCTA 601
| | | | | | | | | | | | | | | | | | | | | |
DB 2213 AAGTAACATGAAAGAAAAGAAAGAAAGTAACACATGAAAGAAAAGAAAGTAAGTA 2272
| | | | | | | | | | | | | | | | | | | | | |
QY 602 AAGAAAAGTTGCAAGACAGCAAAAGATTTAGAACAAAGAAAGCTGATACGAAAAAA 661
| | | | | | | | | | | | | | | | | | | | | |
DB 2273 TACATGAAAGAAAAGAAAAGAAAGTAATCATGAAAGAAAAGAAAGTAACATG 2332
| | | | | | | | | | | | | | | | | | | | | |
QY 662 ATTTGAAAGAAAAGAAAGCAATGAGATATATTGCAAGAGATTTATATGTCGTTAG 721
| | | | | | | | | | | | | | | | | | | | | |
DB 2333 AAGAAAGAAAAGAAAGAAAGTAACACATGAAAGAAAAGAAAGTAACATGAAAG 2392
| | | | | | | | | | | | | | | | | | | | | |
QY 722 AAATACAGCATATGAACCTCCATCAGAAAATGAAACGCGATATATATATACCATCAAT 781
| | | | | | | | | | | | | | | | | | | | | |
DB 2393 AAAAAAGAAAGTAACATGAAAGAAAGAAAGAAAGTAACATGAAAGAAAGAAAGTA 2452
| | | | | | | | | | | | | | | | | | | | | |
QY 782 CTTCTTAACTCAGAGCAACAGAGGAATGAGATTCCAGAAAATATCTTAAATAG 841
| | | | | | | | | | | | | | | | | | | | | |
DB 2453 AAGTAACATGAAAGAAAAGAAAGAAAGTAACATGAAAGAAAGAAAGTAAGTA 2512
| | | | | | | | | | | | | | | | | | | | | |
QY 842 AAAAAACAATAGGAATCTATTCAACAAATGTTGGAAGCAAGGATATCAATAAG 901
| | | | | | | | | | | | | | | | | | | | | |
DB 2513 CACATGAAAGAAAGAAAAGTAACATGAAAGAAAGAAAGTAACATGAAAG 2572
| | | | | | | | | | | | | | | | | | | | | |
QY 902 GACATCTTGAAGAAAGAAAGATGCTTCAATTAACAGAACAAAGAAAGATTAATCTG 961
| | | | | | | | | | | | | | | | | | | | | |
DB 2573 AAGAAAAGTAACATATGAAAGAAAGAAAGAAAGAAAGTAACATGAAAG 2632
| | | | | | | | | | | | | | | | | | | | | |
QY 962 CTGACATACAAATATCATATATAGAGACAGTAATATTTCTGATGTTAATGATTTTCAA 1021
| | | | | | | | | | | | | | | | | | | | | |
DB 2633 AAGAAAAGTAACATGAAAGAAAGAAAGAAAGTAACATGAAAGAAAGAAAGTA 2692
| | | | | | | | | | | | | | | | | | | | | |
QY 1022 TAAGTAAGTATAGC---ATGAAATTAAGTCTGAATATGACGATCTTAATAGATGAG 1078
| | | | | | | | | | | | | | | | | | | | | |
DB 2693 TACATGAAAGAAAGAAAGAAAGAAAGTAAGTGAAGAAAGAAAGAAAGAAAG 2752
| | | | | | | | | | | | | | | | | | | | | |
QY 1079 AAGAAAGTATGAAAGACTTGAAGCAATTTAAGCCTAATTTGCAATATGACAATTTCCAAG 1138
| | | | | | | | | | | | | | | | | | | | | |
DB 2753 AAGAAAGAAAGAAAGAAAGATGAGAGAGAAAGAAAGAAAGAAAGATGAGAGAG 2812
| | | | | | | | | | | | | | | | | | | | | |
QY 1139 ATGAAGAAAACATATGAAATTTATAAAGACTTGAAGATTTGATGAGAAAATGAATTT 1198
| | | | | | | | | | | | | | | | | | | | | |
DB 2813 AAGAAAGAAAGAAAGAAATGAGAGAGAAAGAAAGAAAGAAATGAAGAAAGAG 2872
| | | | | | | | | | | | | | | | | | | | | |
QY 1199 TAGATGATTTAGATGAAGAAATGAAATATCATGAAAGAAATTTCTGAAGAAAATTA 1258
| | | | | | | | | | | | | | | | | | | | | |
DB 2873 AAGAAAAGAAAGCAATGAAAGAAAGTAACATGAAAGAAAGAAAGAAAGTAACAC 2932
| | | | | | | | | | | | | | | | | | | | | |
QY 1259 AAAAAAGAAAGAAATATGAAAGAAAGAAAGTAATTTTAAACCAATGATAAAGTT 1318
| | | | | | | | | | | | | | | | | | | | | |

CC monitoring RNaseH activity in real time. The methods are also useful for
CC monitoring enzymatic degradation of an RNA-DNA duplex. In an example of
CC the invention, the activity of the RNaseH from E. coli or HIV reverse
CC transcriptase is measure by the method of the invention. In order to test
CC the specificity of the RNaseH cleavage of the target-probe complex, a DNA
CC or RNA contaminant is added to the assay reaction. This sequence
CC corresponds to the single stranded homopolymeric polyA RNA contaminant
CC for testing this activity.

XX
XX Sequence 1000 BP; 1000 A; 0 C; 0 G; 0 T; 0 U; 0 Other;

Query Match 12.2%; Score 180.2; DB 12; Length 1000;

Best Local Similarity 48.8%; Pred. No. 5.3e-17;

Matches 485; Conservative 0; Mismatches 508; Indels 0; Gaps 0;

```
Qy 397 AAGAGAAAGTTGCAAGACAAAGCGATTAGACAGAGAGACTTGTAAGAAAG 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60

Qy 457 TTGCAGACACAAAGCGATTGAAACAGAGAGCGTCTAAAGAAAGTTGCAAG 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 120

Qy 517 CAACAAAGCGATTAGAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGACAAAGC 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 180

Qy 577 GATTAGACACAGAGAGCGTCTAAAGAAAGTTGCAAGACAGAGAGATTAGAA 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 240

Qy 637 CAAGAGAGCGTGTACGAAAAAAAAATTAGAAAGAAAGAACTGAGATATTTA 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 300

Qy 697 GCAGAGATTATATGTGCTTGAATAACACGCTATAGAACTCCATCAGAAATGAA 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 360

Qy 757 CGTGATTTATATACCAATCTTCTTACCTCAGACAAAGAGGATAGTAGA 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 420

Qy 817 GATTCCAGAGAAATATCTATATAGAAAAACAATGAGAACTTATTAACAATGTT 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 480

Qy 877 GAAGAGCAAGGATATACATTAAGGACATCTTGAAAGAAAGAAAGTGTCAATAAA 936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 540

Qy 937 CCAGAACCAAAAGAGATTAATCTGCTGACATCAATCATTAATTAGACAGTAAT 996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 600

Qy 997 ATTTCTGATGTTAATGATTTCAATAAGTAAGTAGAGATGAATTAAGTCTGAATAT 1056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 660

Qy 1057 GACGATTTATATATGATGAAGAGAGATGTAAGCTTAGAGCAATTTAAAGCTATT 1116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 720

Qy 1117 GTGCATATGACAAATTTCCAGATGAGAAACAATAGAAATTTTAAAGACTAGAGAT 1176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 721 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 780

Qy 1177 TTGATAGAGAAATGAAATTTTATGATGATTTAGATGAAGAAATTCATCAGAA 1236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 781 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 840

Qy 1237 GAATTTATCTGAAGAAATTAAGAAAGAAATGTAAGAAAGAAAGATTAATAT 1296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 841 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 900

Qy 1297 TTTAAACCAATATATATTTGTATGATGACATTTTAAATATTAATAATGATTAAG 1356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 901 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 960

Qy 1357 CAGGTTAATGAGAAAGGAAATTCATTAATA 1389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 961 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 993
```

RESULT 12

AD062832/c

ID AD062832 standard; RNA; 1000 BP.

XX AD062832;

XX 07-OCT-2004 (first entry)

XX Homopoly-U contaminant for RNaseH activity assay.

XX ss; nuclease-mediated cleavage; target nucleic acid; RNaseH activity;

KM fluorophore; fluorescence quencher; fluorescent signal; HIV;

KM reverse transcriptase.

OS Synthetic.

XX WO2004059012-A1.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003MO-US040879.

XX 23-DEC-2002; 2002US-0436125P.

XX (AMHP) WYETH.

XX PI Olson MM, O'Connell JF;

XX WPI: 2004-543471/52.

XX Detecting a nuclease-mediated cleavage of a target nucleic acid, useful

XX for detecting and monitoring RNase H activity, comprises hybridizing a

XX target nucleic acid to a fluorescently labeled oligonucleotide probe.

XX Example 1; SEQ ID NO 3; 61pp; English.

CC The invention relates to a method of detecting a nuclease-mediated
CC cleavage of a target nucleic acid or measuring a RNaseH activity of an
CC agent by hybridizing a target nucleic acid to a fluorescently labelled
CC oligonucleotide probe complementary to the target nucleic acid and
CC containing a fluorophore at one terminus and a quenching group at the
CC other terminus and contacting the probe-target hybrid with an agent
CC having nuclease activity. When the oligonucleotide probe is unhybridised
CC to the target nucleic acid, the probe adopts a conformation that places
CC the fluorophore and quencher in such proximity that the quencher quenches
CC the fluorescent signal of the fluorophore. Formation of the probe-target
CC hybrid causes sufficient separation of the fluorophore and quencher to
CC reduce quenching of the fluorescent signal of the fluorophore followed by
CC contacting the probe-target hybrid with an agent having nuclease activity
CC to selectively cleave the target nucleic acid and thus release the intact
CC probe. Detection of the release of the probe is by measuring a decrease
CC in the fluorescent signal of the fluorophore as compared to the signal of
CC the probe-target hybrid. The methods are useful for detecting and
CC monitoring RNaseH activity in real time. The methods are also useful for
CC monitoring enzymatic degradation of an RNA-DNA duplex. In an example of
CC the invention, the activity of the RNaseH from E. coli or HIV reverse
CC transcriptase is measure by the method of the invention. In order to test
CC the specificity of the RNaseH cleavage of the target-probe complex, a DNA
CC or RNA contaminant is added to the assay reaction. This sequence
CC corresponds to the single stranded homopolymeric polyU RNA contaminant
CC for testing this activity.

Sequence 1000 BP; 0 A; 0 C; 0 G; 0 T; 1000 U; 0 Other;

```
Query Match 12.2%; Score 180.2; DB 12; Length 1000;
Beet Local Similarity 48.8%; Pred. No. 5.3e-17;
Matches 485; Conservative 0; Mismatches 508; Indels 0; Gaps 0;

QY AAAAGAAAGTTCCAGAACCAACAAACGATTAGAAACAAGAGAGACTTGCTAAAGAAAAG 456
DB AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 941
QY 1000 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
DB 457 TTGCAAGAACCAACGATTAGAAACAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAA 516
QY 940 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 881
DB 517 CAACAAGGATTTAGAACAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACCAAGC 576
DB 880 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 821
QY 577 GATTAGAACAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAGCAGCAAGAGATTAGAA 636
DB 820 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 761
QY 637 CAAGAGAGCTGATACGAAAAAAAAATTAGAAAGAAAAGCAATGAGATATATTA 696
DB 760 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 701
QY 697 GCAGAGGATTTATATGCTGTTAGAAATACAGCTATAGAACTTCATCAGAAAATGAA 756
DB 700 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 641
QY 757 CGTGATATATATACACATCAATCTCTTAACTCAGAGCAACAGAGGATATAGAA 816
DB 640 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 581
QY 817 GATTCAGAGAAATATCTATATATAGAAAAACAATAGAAATCTATTAACAATGTT 876
DB 580 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 521
QY 877 GAAGACGAGGATTTACATTAAGACATCTTGAGAGAAAGAGATGTTCAATTA 936
DB 520 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 461
QY 937 CCAGAACCAAAAGAGATTAATCTGCTGACATACAAATCATACATTAGAGACAGTAAT 996
DB 460 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 401
QY 997 ATTTCTGATGTTAATGATTTTCAAAATAGTAAGATGAGATGAATTAAGCTGATAT 1056
DB 400 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 341
QY 1057 GACGATTCATTAATAGATGAAGAGAGATGAGAGACTTACGAAATTTAAGCTATT 1116
DB 340 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 281
QY 1117 GTGCAATATGACAAATTTCCAGATGAAGAAACATAGGAATTTATTAAGATAGAT 1176
DB 280 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 221
QY 1177 TTGATAGAGAAATAGAAATTTAGATGATTAAGAGAAATAGAAATATCATCAGAA 1236
DB 220 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 161
QY 1237 GAATTTCTGAGAGAAAAATTAAGAAAGAGAAATATGAAAGAAACAAAGATTAAT 1296
DB 160 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 101
QY 1297 TTTAAACCAATGATTAAGATTTGATGATGACATATTAATAATATATATGATAAG 1356
DB 100 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 41
QY 1357 CAGGTTAATAGAGAAAGAGAAATTCATTA 1389
DB 40 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8
```

RESULT 13

AA087587
ID AA087587 standard; DNA; 1686 BP.

XX AA087587;

AC 27-AUG-2003 (revised)

DT 19-DEC-1995 (first entry)

XX DNA encoding Leucocytozoan protozoan structural protein epitope.

DE Leucocytozoan protozoa; structural protein; epitope; vaccine; fowl;

KW Leucocytozoanosis; treatment; ss.

OS Leucocytozoan.

XX JP07089995-A.

XX 04-APR-1995.

XX 10-SEP-1993; 93JP-00226078.

XX 10-SEP-1993; 93JP-00226078.

XX (DOBU-) DOBUTSUYO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.

PA (NISS-) NISSEIKEN KK.

DR WPI; 1995-167252/22.

XX P-PsDB; AAR70491.

PT Immune inducing polypeptide against Leucocytozoan protozoa - useful in

PS production of vaccines for treatment of Leucocytozoanosis in fowl.

XX Claim 1; Page 12-14; 20pp; Japanese.

CC AA087587-69 encode polypeptides having a whole or partial epitope of a

CC structural protein of Leucocytozoan protozoa (see AAR70491-93). The

CC polypeptides and DNA encoding them are useful in the production of

CC vaccines for the treatment of Leucocytozoanosis of fowl. (Updated on 27-

CC AUG-2003 to correct OS field.)

XX Sequence 1686 BP; 915 A; 129 C; 382 G; 260 T; 0 U; 0 Other;

Query Match 12.1%; Score 179.4; DB 2; Length 1686;

Beet Local Similarity 46.9%; Pred. No. 6.6e-17;

Matches 595; Conservative 0; Mismatches 671; Indels 3; Gaps 1;

QY 2 AAGAACCAAAAGCATCTAGAACAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAC 61

DB 101 AAGAGAAATGAGAAATGAAATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 160

QY 62 AAAGGATTTAGAACAGATAGACTTGTAAAGAAAAGTTACAGAGCAGAAAGCATT 121

DB 161 AACAGAACCAAGAAATCGAGAAAGAAAGAAAGAAAGATGAGAGAAACAAAGAGAG 220

QY 122 TAGAACAGAGAGACTTGTAAAGAAAAGTTGCAAGAACCAAGCCGATCTAGAACAG 181

DB 221 ATGAAGAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 280

QY 182 AGAGAGCTGCTAAAGAAAAGTTGCAAGAACCAAGAGCAGATTTAGAACAGAGAGAG 241

DB 281 AAG 340

QY 242 CTAAAGAAAAGTTGCAAGAACCAAGAGCAGATTTAGAACAGATGACTTCTAAAGAA 301

DB 341 ATGAAG 400

QY 302 AGTTACAG 361

DB 401 AAG 460


```
Qy 362 AACACAAAGCGATTTAGAACAGAGAGACGCTGTAAGAAAGTTGCAAGAACACAA 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 461 AAGAACAAACAGAACAGATGAAAGAGAACAGATGAAGAAACAGATGAAGAGAG 520
Qy 422 GCGATTTAGAACAGAGAGACTTCTAAGAAAGTTGCAAGAACACAAAGCGATTTAG 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 521 AAGAACAAAGAACAGAGAGATGAAAGAGAACAGAGAGATGAAAGAGAGAGAG 580
Qy 482 AACAGAGAGAGCGTCTAAGAAAGTTGCAAGAACACAAAGCGATTTAGAACAGAG 541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 581 ATGAAGATGAAGAAAGAAAGAAAGAAAGAAAGAAAGATGAAGAAAGATATATG 640
Qy 542 GACCTGCTAAGAAAGTTGCAAGAACACAAAGCGATTTAGAACAGAGAGAGTCTA 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 641 AAGAAAGAAAGAGAGAGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 700
Qy 602 AAGAAAGTTGCAAGAGAGAGAGATTTAGAACAAAGAGAGAGAGAGAGAGAGAG 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 701 AAGAAAGAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 757
Qy 662 ATTTAGAAAGAAAGAGAGAGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 758 CAAAGAAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 817
Qy 722 AAGATCAAGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 818 AAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 877
Qy 782 CTTCTTACCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 878 CAAATGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 937
Qy 842 AAAAAACAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 938 TAAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 997
Qy 902 GACATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 961
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 998 ATGAATCAAAATCCGAAATGAATTCATCAATTAATTTGAAGAGAGAGAGAG 1057
Qy 962 CTGACATGACAAATCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1058 CATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1117
Qy 1022 TAACTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1118 TTGAAGAGAGTTGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1177
Qy 1082 AAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1178 TAGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
Qy 1142 AAGAAACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1238 CACCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1297
Qy 1202 ATGATTTAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1298 AAGAAACATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
Qy 1262 AAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1358 TAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 14
AAAT70099
ID AAAT70099 standard; DNA; 3579 BP.
AC AAAT70099;
XX
XX 07-NOV-2000 (first entry)
```

```
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:232.
XX
XX Plasmodium falciparum; chromosome 2, human malaria parasite; vaccine;
KW antimalarial; malaria; protozoa; infection; insecticide; ds.
XX
XX Plasmodium falciparum.
PN M0200025728-A2.
XX
XX 11-MAY-2000.
PD
XX
XX 05-NOV-1999; 99MO-US026796.
PF
XX
XX 05-NOV-1998; 98US-0107131P.
PR
XX
XX (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
XX Hoffman S, Carucci D, Gardner M, Venter JC;
PI WPI; 2000-365347/31.
DR
XX
XX The present invention describes proteins and their fragments (I) encoded
PT by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
PT Also described are: (I) nucleotide sequences (I) encoding (I); and (2)
PT vaccines against P. falciparum infection comprising (I) or (II). (I) and
PT (II) are useful for the development of vaccines against P. falciparum
PT infection. (I) and polyclonal antisera or a monoclonal antibody raised to
PT immunogens comprising the sequences of (I), are useful in the detection
PT of infection with P. falciparum. Furthermore, (I) (especially when they
PT are rIFNs or secreted or membrane proteins) can aid the identification
PT of drugs to treat or prevent P. falciparum infection, or they can be used
PT to identify drug resistance in P. falciparum. Sequencing of the
PT Plasmodium chromosome 2 and the subsequent identification of proteins
PT encoded by it will help to expand our understanding of parasite biology,
PT a process hampered by the complexity of the parasite life cycle, and
PT provide new targets for vaccine and drug development. Parasite resistance
PT to drugs and mosquito resistance to insecticides have led to a resurgence
PT of malaria in many parts of the world, and there is a pressing need for
PT vaccines and new drugs. AAAT70078 to AAAT70287 and AAB18144 to AAB18352
PT represent nucleotide and protein sequences given in the present
PT invention, but which are not specifically mentioned within the
PT specification
XX
XX
XX Sequence 3579 BP; 1904 A; 398 C; 552 G; 725 T; 0 U; 0 Other;
SO
Query Match 11.9%; Score 176.6; DB 3; Length 3579;
Beet Local Similarity 46.3%; Pred. No. 1.6e-16;
Matches 661; Conservative 0; Mismatches 759; Indels 9; Gaps 2;
Qy 36 TGCCTAAGAAAGTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 95
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 744 TGATTAATACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 96 AAGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 804 TAGAAGACATTTCTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 156 AGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 864 AATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 923
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 216 AAGCATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 924 TGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 983
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 23:09:13 ; Search time 7825 Seconds
(without alignments)
10765.750 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 1482

Sequence: 1 CAAGACACAAACAAAGCGATCT.....ATAATTTTATGAAACTATTA 1482

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pt:*
9: gb_to:*
10: gb_gst:*
11: gb_ay:*
12: gb_un:*
13: gb_vt:*
14: gb_hcg:*
15: gb_d1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1482	100.0	1482	6	ARI62970 Sequence
2	1482	100.0	1482	6	ARI78259 Sequence
3	1482	100.0	1496	2	PFLSA13
4	1480.4	99.9	1482	6	ARI62974 Sequence
5	1480.4	99.9	1482	6	ARI78263 Sequence
6	1446	97.6	1493	6	A28743
7	1446	97.6	1493	6	ARI62967 Sequence
8	1446	97.6	1493	6	ARI78256 Sequence
9	1442	97.3	5970	2	PFLSA1G
10	1406.8	94.9	253001	2	AE014834
11	950	64.1	950	6	A28740
12	950	64.1	950	6	ARI62961
13	950	64.1	950	6	ARI78250
14	843	56.9	843	2	PFLSA1C
15	841.4	56.8	843	2	PFLSA1B
16	841.4	56.8	843	2	PFLSA1D
17	841.4	56.8	843	2	PFLSA1H
18	841.4	56.8	843	2	PFLSA1I

19	839.8	56.7	843	2	PFLSA1E	L40887 Plasmodium
20	839.8	56.7	843	2	PFLSA1F	L40888 Plasmodium
21	839.8	56.7	843	2	PFLSA1G	L40893 Plasmodium
22	839.8	56.7	843	2	PFLSA1H	L40834 Plasmodium
23	838.2	56.6	843	2	PFLSA1M	L40835 Plasmodium
24	836.6	56.5	843	2	PFLSA1N	L40947 Plasmodium
25	836.6	56.5	843	2	PFLSA1D	L40886 Plasmodium
26	836.6	56.5	843	2	PFLSA1N	L40836 Plasmodium
27	836.6	56.5	843	2	PFLSA1O	L40837 Plasmodium
28	835	56.3	843	2	PFLSA1C	L40910 Plasmodium
29	835	56.3	843	2	PFLSA1J	L40892 Plasmodium
30	833.4	56.2	843	2	PFLSA1A	L40908 Plasmodium
31	833.4	56.2	843	2	PFLSA1B	L40909 Plasmodium
32	791.8	53.4	795	2	AF246996	AF246996 Plasmodium
33	503.4	34.0	1368	11	AY751501	AY751501 Synthetic
34	472.8	31.9	988	2	PFLSA1S	Z30319 P. falciparum
35	472.8	31.9	988	6	A28742	A28742 cDNA for LS
36	472.8	31.9	988	6	ARI62963	ARI62963 Sequence
37	472.8	31.9	988	6	ARI78252	ARI78252 Sequence
38	470.8	31.8	954	6	ARI62966	ARI62966 Sequence
39	470.8	31.8	954	6	ARI78255	ARI78255 Sequence
40	280.6	18.9	198224	9	AC149597	AC149597 Mus muscu
41	274.4	18.5	208332	9	AC122227	AC122227 Mus muscu
42	252.4	17.0	172716	14	AC115383	AC115383 Rattus no
43	250.8	16.9	184015	9	AC144849	AC144849 Mus muscu
44	237.4	16.0	161482	14	AC165336	AC165336 Mus muscu
45	225.8	15.2	2069	6	E10125	E10125 DNA encodin

ALIGNMENTS

RESULT 1
LOCUS ARI62970 1482 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 41 from patent US 6270771.
ACCESSION ARI62970
VERSION ARI62970.1 GI:16233435
KEYWORDS
ORGANISM Unknown.
SOURCE Unknown.
REFERENCE 1 (bases 1 to 1482)
AUTHORS Guerlin-Marchand,C. and Druilhe,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6270771-A 41 07-AUG-2001;
FEATURES
source Location/Qualifiers
1..1482
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 1482; DB 6; Length 1482;
Best Local Similarity 100.0%; Pred. No. 1,3e-158;
Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGACACAAACGATCTTGAACAAGAGAGAGCTCTTAAGAAAGTTGCAACAA 60
DB 1 CAAGACACAAACGATCTTGAACAAGAGAGAGCTCTTAAGAAAGTTGCAACAA 60
QY 61 CAAGCGATTTAGAAACAGATAGCTTCTTAAGAAAGTTGCAAGAGCGCAAGCGAT 120
DB 61 CAAGCGATTTAGAAACAGATAGCTTCTTAAGAAAGTTGCAAGAGCGCAAGCGAT 120
QY 121 TTAGACACAGAGAGACTTGTCTTAAGAAAGTTGCAAGACCAAGCGATCTTGAACAA 180
DB 121 TTAGACACAGAGAGACTTGTCTTAAGAAAGTTGCAAGACCAAGCGATCTTGAACAA 180
QY 181 GAGAGAGCTGTCTTAAGAAAGTTGCAAGACCAAGCGATTTTGAACAGAGAGAGCT 240
DB 181 GAGAGAGCTGTCTTAAGAAAGTTGCAAGACCAAGCGATTTTGAACAGAGAGAGCT 240

QY 241 GCTAAGAAAAGTTCAGAGAACAAAGCGATTTAGAACAGATGACTTCTTAAGAA 300
DB 241 GCTAAGAAAAGTTCAGAGAACAAAGCGATTTAGAACAGATGACTTCTTAAGAA 300
QY 301 AAGTTACAAAGACGAGCAAGCGATTTAGAACAGAGACGCTCTTAAGAAAAGTTGCA 360
DB 301 AAGTTACAAAGACGAGCAAGCGATTTAGAACAGAGACGCTCTTAAGAAAAGTTGCA 360
QY 361 GAACAAACAAGCGATTTAGAACAGAGACGCTCTTAAGAAAAGTTGCAAGAACAA 420
DB 361 GAACAAACAAGCGATTTAGAACAGAGACGCTCTTAAGAAAAGTTGCAAGAACAA 420
QY 421 AGCGATTTAGAACAGAGACGCTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
DB 421 AGCGATTTAGAACAGAGACGCTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
QY 481 GAACAAAGAGACGCTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 540
DB 481 GAACAAAGAGACGCTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 540
QY 541 AGACGCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGCTCT 600
DB 541 AGACGCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGCTCT 600
QY 601 AAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGAGCTGATCGAAAAA 660
DB 601 AAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGAGCTGATCGAAAAA 660
QY 661 AATTTAGAAAAGAAAAGAACATGAGATATATTAGCAGAGATTTATATGCTGTTA 720
DB 661 AATTTAGAAAAGAAAAGAACATGAGATATATTAGCAGAGATTTATATGCTGTTA 720
QY 721 GAAATACCAAGCTATGAACTTCATCAGAAAATGAAACGAGATTTATATACACATCA 780
DB 721 GAAATACCAAGCTATGAACTTCATCAGAAAATGAAACGAGATTTATATACACATCA 780
QY 781 TCTTCTTTACCTCAGAGCAACAGAGGAAATAGTAGAGATTTCCAGAAAATCTATAA 840
DB 781 TCTTCTTTACCTCAGAGCAACAGAGGAAATAGTAGAGATTTCCAGAAAATCTATAA 840
QY 841 GAAAAAACAATAGAGATCTATTACACAAATGTTGAGACGAGAGATTTACATAA 900
DB 841 GAAAAAACAATAGAGATCTATTACACAAATGTTGAGAGAGAGATTTACATAA 900
QY 901 GGAACATCTTGAAGAAAAGAAAGTGTGTAATTAACCAAGCAAAAAGAAATTAAT 960
DB 901 GGAACATCTTGAAGAAAAGAAAGTGTGTAATTAACCAAGCAAAAAGAAATTAAT 960
QY 961 GCTGACATACAAATCATCATTTAGAGACGATTAATATTCTGATGTTAATGATTTCA 1020
DB 961 GCTGACATACAAATCATCATTTAGAGACGATTAATATTCTGATGTTAATGATTTCA 1020
QY 1021 ATTAAGTATAGAGATGAAATTAAGTGTGTAATTAAGATTTCAATTAATAGTAGAA 1080
DB 1021 ATTAAGTATAGAGATGAAATTAAGTGTGTAATTAAGATTTCAATTAATAGTAGAA 1080
QY 1081 GAAAGTATAGAGACCTTAAGCAATTTAAGCCTAATGTCATTAAGACAAATTTCAAG 1140
DB 1081 GAAAGTATAGAGACCTTAAGCAATTTAAGCCTAATGTCATTAAGACAAATTTCAAG 1140
QY 1141 GAAGAAAACATAGAAATTTAAGAACTAGAAAGATTTAGAGAGAAAATGAAATTTA 1200
DB 1141 GAAGAAAACATAGAAATTTAAGAACTAGAAAGATTTAGAGAGAAAATGAAATTTA 1200
QY 1201 GATGATTTAGAGAGAAATAGAAAATCATCAAGAAATTTATCTGAGAAAATTAATA 1260
DB 1201 GATGATTTAGAGAGAAATAGAAAATCATCAAGAAATTTATCTGAGAAAATTAATA 1260
QY 1261 AAGAGAAAAGTATAGAAAACAAAGATTAATTTAAGCAAAATGATTAAGTTG 1320
DB 1261 AAGAGAAAAGTATAGAAAACAAAGATTAATTTAAGCAAAATGATTAAGTTG 1320
QY 1321 TATGATGACATATTTAAAAATATAAAAATGATTAAGCAGTTAATTAAGAAAAGAA 1380

DB 1321 TATGATGACATATTTAAAAATATAAAAATGATTAAGCAGTTAATTAAGAAAAGAA 1380
QY 1381 TTCAATTAAGATGTTTCATATATTTGACGAGACATGAAATTTTACAGATCGTGAT 1440
DB 1381 TTCAATTAAGATGTTTCATATATTTGACGAGACATGAAATTTTACAGATCGTGAT 1440
QY 1441 GAGTTATCTGAAGATATTAATTAATTTAATTTAAGAACTATTA 1482
DB 1441 GAGTTATCTGAAGATATTAATTAATTTAATTTAAGAACTATTA 1482

RESULT 2

AR178259 1482 bp DNA linear PAT 20-APR-2002

LOCUS

Sequence 41 from patent US 6319502.

DEFINITION

AR178259

ACCESSION

AR178259.1

VERSION

GI:20219397

KEYWORDS

Unknown.

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 1482)

AUTHORS

Guerin-Marchand, C. and Druille, P.

TITLE

Peptide sequences specific for the hepatic stages of P. falciparum

JOURNAL

bearing epitopes capable of stimulating the T lymphocytes

FEATURES

Patent: US 6319502-A 41 20-NOV-2001;

ORIGIN

source

1. 1482

/organism="unknown"

/mol_type="unassigned DNA"

Query Match

Best Local Similarity 100.0%; Score 1482; DB 6; Length 1482;

Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGAACAAAGCGATCTAGAACAGAGACGCTGTAAGAAAAGTTGCAAGAACAA 60
DB 1 CAAGAACAAAGCGATCTAGAACAGAGACGCTGTAAGAAAAGTTGCAAGAACAA 60
QY 61 CAAGCGATTTAGAACAGATGACTTCTTAAGAAAAGTTCAAGAGCAGCAAGCGAT 120
DB 61 CAAGCGATTTAGAACAGATGACTTCTTAAGAAAAGTTCAAGAGCAGCAAGCGAT 120
QY 121 TTAGAACAGAGACGCTGTAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
DB 121 TTAGAACAGAGACGCTGTAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
QY 181 GAGAGACGCTGTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGCT 240
DB 181 GAGAGACGCTGTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGCT 240
QY 241 GCTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATGACTTCTTAAGAA 300
DB 241 GCTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATGACTTCTTAAGAA 300
QY 301 AAGTTACAAAGACGAGCAAGCGATTTAGAACAGAGACGCTCTTAAGAAAAGTTGCA 360
DB 301 AAGTTACAAAGACGAGCAAGCGATTTAGAACAGAGACGCTCTTAAGAAAAGTTGCA 360
QY 361 GAACAAACAAGCGATTTAGAACAGAGACGCTGTAAGAAAAGTTGCAAGAACAA 420
DB 361 GAACAAACAAGCGATTTAGAACAGAGACGCTGTAAGAAAAGTTGCAAGAACAA 420
QY 421 AGCGATTTAGAACAGAGACGCTGTAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
DB 421 AGCGATTTAGAACAGAGACGCTGTAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
QY 481 GAACAAAGAGACGCTGTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 540
DB 481 GAACAAAGAGACGCTGTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 540

Oy	541	GGAGCGCTAAAGAAAGAGCTTCAGAAACAAGAAACGCTTTATGAAACAAGAGACGCGCT	600
Db	541	AGACGCTGCTAAAGAAAGAGCTTCAGAAACAAGAAACGCTTTATGAAACAAGAGACGCGCT	600
Oy	601	AAAGAAAGAGTTGCAGAGACAGCAAGAGATTAGAACCAAGAGAGCGCTGATACCAAAAA	660
Db	601	AAAGAAAGAGTTGCAGAGACAGCAAGAGATTAGAACCAAGAGAGCGCTGATACCAAAAA	660
Oy	661	AATTTAGAAAGAAAAAGAAACATGAGATATATTAGCAGAGATTATATGCTGCTTTA	720
Db	661	AATTTAGAAAGAAAAAGAAACATGAGATATATTAGCAGAGATTATATGCTGCTTTA	720
Oy	721	GAAATACAGCATATGAACTTCATCAGAAAAATGAACGTGATATTTATATCCATCA	780
Db	721	GAAATACAGCATATGAACTTCATCAGAAAAATGAACGTGATATTTATATCCATCA	780
Oy	781	TCTTCTTACCTCAGAGACAGAGGGAATAGTAGAGATTCCAGAGAAATATCTATATA	840
Db	781	TCTTCTTACCTCAGAGACAGAGGGAATAGTAGAGATTCCAGAGAAATATCTATATA	840
Oy	841	GAAGAAACAAATAGAGAACTTATTAACAATAATGTTGAAGACGAGGGATATACATAA	900
Db	841	GAAGAAACAAATAGAGAACTTATTAACAATAATGTTGAAGACGAGGGATATACATAA	900
Oy	901	GGACATCTTGAAAGAAAGAAAGATGCTTCATTAAGAACAGAGAACAAAAAGAGATTAATCT	960
Db	901	GGACATCTTGAAAGAAAGAAAGATGCTTCATTAAGAACAGAGAACAAAAAGAGATTAATCT	960
Oy	961	GCTGACATACAAAAATCATACATTAGAGACAGTAATATTTCTGATGTTAATGATTTTCAA	1020
Db	961	GCTGACATACAAAAATCATACATTAGAGACAGTAATATTTCTGATGTTAATGATTTTCAA	1020
Oy	1021	ATTAAGTAAGTATGAGATGAATAAATGCTGATATGACATTCATTATATGATGAAGAA	1080
Db	1021	ATTAAGTAAGTATGAGATGAATAAATGCTGATATGACATTCATTATATGATGAAGAA	1080
Oy	1081	GAAAGTATGAAAGACTTGAAGAAATTTAAGCCTATTTGCAATATGACAAATTTCCAAAT	1140
Db	1081	GAAAGTATGAAAGACTTGAAGAAATTTAAGCCTATTTGCAATATGACAAATTTCCAAAT	1140
Oy	1141	GAAGAAAAATAGAAATTTATTAAGACATAGAGATTTGATAGAGAAAAATGAAAAATTTA	1200
Db	1141	GAAGAAAAATAGAAATTTATTAAGACATAGAGATTTGATAGAGAAAAATGAAAAATTTA	1200
Oy	1201	GATGATTTATGATGAAAGAAATGAAATAATCATACAGAAATTTCTGAAAGAAAAATTA	1260
Db	1201	GATGATTTATGATGAAAGAAATGAAATAATCATACAGAAATTTCTGAAAGAAAAATTA	1260
Oy	1261	AAAGAAAGAAATATGAAAAAAACAAGATATAATTTTAAACCAAAATGATTAAGAAATTTG	1320
Db	1261	AAAGAAAGAAATATGAAAAAAACAAGATATAATTTTAAACCAAAATGATTAAGAAATTTG	1320
Oy	1321	TATGATGAGCATATTTAABAAAATATBAABATGATTAAGCAGGTTAATAGAGAAAAAGAAAA	1380
Db	1321	TATGATGAGCATATTTAABAAAATATBAABATGATTAAGCAGGTTAATAGAGAAAAAGAAAA	1380
Oy	1381	TTTCATTAATCATTTGTTTCATATATTGACGAGACAAATGAATTTTACAGATCGTGAT	1440
Db	1381	TTTCATTAATCATTTGTTTCATATATTGACGAGACAAATGAATTTTACAGATCGTGAT	1440
Oy	1441	GAGTTATCTGAGATATTAATAATTTTATGAAACATATA	1482
Db	1441	GAGTTATCTGAGATATTAATAATTTTATGAAACATATA	1482

	RESULTS	REFERENCE
LOCUS	PFLSA13	
DEFINITION	P.falciparum LSA-1 gene for liver stage antigen-1 (3').	
ACCESSION	Z30320	
VERSION	Z30320.1 GI:510183	
KEYWORDS	liver stage antigen-1; LSA-1 gene.	
SOURCE	Plasmodium falciparum (malaria parasite <i>P. falciparum</i>)	
FEATURES		
ORIGIN		
COMMENTS		

[illegible]

Qy	181	GAGAGCGTCTTAAGAAAAAGTTGCAAGAACAA	CAAAAGCGATTTTAGAACAAAGAGACGT	240
Db	181	GAGAGCGTCTTAAGAAAAAGTTGCAAGAACAA	CAAAAGCGATTTTAGAACAAAGAGACGT	240
Qy	241	GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAA	CAAGATAGACTTGGCTAAAGAA	300
Db	241	GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAA	CAAGATAGACTTGGCTAAAGAA	300
Qy	301	AAGTTACAAAGACGACAAAGCGATTTAGAACAAAGAGAC	GTGCTAAAGAAAAGTTGCAA	360
Db	301	AAGTTACAAAGACGACAAAGCGATTTAGAACAAAGAGAG	GTGCTAAAGAAAAGTTGCAA	360
Qy	361	GAAACAACAAAGCGATTTAGAACAAAGAGACGTGCTTA	AAAGAAAAGTTGCAAGAACAA	420
Db	361	GAAACAACAAAGCGATTTAGAACAAAGAGACGTGCTTA	AAAGAAAAGTTGCAAGAACAA	420
Qy	421	AGCGATTTAGAACAAAGAGACTTGGCTTAAGAAAAGTTGCA	AGAACAAAGCGATTTA	480
Db	421	AGCGATTTAGAACAAAGAGACTTGGCTTAAGAAAAGTTGCA	AGAACAAAGCGATTTA	480
Qy	481	GAAACAAGAGACGTGCTTAAGAAAAGTTGCAAGAACAA	CAAGCGATTTAGAACAGAG	540
Db	481	GAAACAAGAGACGTGCTTAAGAAAAGTTGCAAGAACAA	CAAGCGATTTAGAACAGAG	540
Qy	541	AGACGTGCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAA	CAAGAGACGTGCT	600
Db	541	AGACGTGCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAA	CAAGAGACGTGCT	600
Qy	601	AAAGAAAAGTTGCAAGAGCAAGAGATTTAGAACAAAGGA	AGCTGATTCGAAAAA	660
Db	601	AAAGAAAAGTTGCAAGAGCAAGAGATTTAGAACAAAGGA	AGCTGATTCGAAAAA	660
Qy	661	AATTTAGAAAAGAAAAGAACATGAGATATATTAGACAGAG	ATTTATAGTCGTTTA	720
Db	661	AATTTAGAAAAGAAAAGAACATGAGATATATTAGACAGAG	ATTTATAGTCGTTTA	720
Qy	721	GAAATACACAGCTTATAGAACTTCCATCAAGAAAATGAA	ACGTGATTTTATATACCATCAA	780
Db	721	GAAATACACAGCTTATAGAACTTCCATCAAGAAAATGAA	ACGTGATTTTATATACCATCAA	780
Qy	781	TCCTCTTACCTGAGCAACAGAGGGAATATGATGAAATTC	CAAGGAAATATCTATATATA	840
Db	781	TCCTCTTACCTGAGCAACAGAGGGAATATGATGAAATTC	CAAGGAAATATCTATATATA	840
Qy	841	GAAAAAAACAATAGAGAAATCTATTACAAACAATGTTGA	AGGACGAAAGGATATACATPAA	900
Db	841	GAAAAAAACAATAGAGAAATCTATTACAAACAATGTTGA	AGGACGAAAGGATATACATPAA	900
Qy	901	GGAATCTTGAAGAAAAGAAAGATGTTCAATTAACCCAG	AAACAAAGAAAGATPAACT	960
Db	901	GGAATCTTGAAGAAAAGAAAGATGTTCAATTAACCCAG	AAACAAAGAAAGATPAACT	960
Qy	961	GCATGACATTAACAAAATCTATTAAGATTAAGAACAG	TAATAATTTCTGATGTTATGATTTCAA	1020
Db	961	GCATGACATTAACAAAATCTATTAAGATTAAGAACAG	TAATAATTTCTGATGTTATGATTTCAA	1020
Qy	1021	ATAAGTATGATGAGATGAAATTAAGTCTGAATATGACGA	TTTCAATTAATAGATGAAGAA	1080
Db	1021	ATAAGTATGATGAGATGAAATTAAGTCTGAATATGACGA	TTTCAATTAATAGATGAAGAA	1080
Qy	1081	GAAATGATGATGAAGACTTGAAGAAATTTAGCGAATTT	AGCGTAATGACAAATTTCCAAAGAT	1140
Db	1081	GAAATGATGATGAAGACTTGAAGAAATTTAGCGAATTT	AGCGTAATGACAAATTTCCAAAGAT	1140
Qy	1141	GAAAGAAAACATAGGAATTTATTAAGAACTAGAAACAT	AGAAAGAAAAGAAATTTTA	1200
Db	1141	GAAAGAAAACATAGGAATTTATTAAGAACTAGAAACAT	AGAAAGAAAAGAAATTTTA	1200
Qy	1201	GATGATTTATGATGAAGAAATGAAAAATTCATCAGAA	GAATATTTCTGAGAAAGAAAATPAAA	1260
Db	1201	GATGATTTATGATGAAGAAATGAAAAATTCATCAGAA	GAATATTTCTGAGAAAGAAAATPAAA	1260
Qy	1261	AAAGAAAAGAAATATGAAAAAACAAGATATATATTTT	AAACCAAAATGATPAAAGTTG	1320

Db	1261	AAAGGAAGAATATGAAAAACAAAGAGTATATATTTTAAACCAATGATTAAGAAGTTTG	1320
Qy	1321	TATGATGACATATTTAAAAATATPAAAAATGATTAAGCAGGTTATTAAGSAAAAAGAAAA	1380
Db	1321	TATGATGACATATTTAAAAATATPAAAAATGATTAAGCAGGTTAATAAGSAAAAAGAAAA	1380
Qy	1381	TTTCATAAATCATTTGTTTCATATATTTTGACGCGAGACATGAAATTTTACAGATCGTGAT	1440
Db	1381	TTTCATAAATCATTTGTTTCATATATTTTGACGCGAGACATGAAATTTTACAGATCGTGAT	1440
Qy	1441	GAGTTATCTGAAGATATACATAATATTTTATGAAACTATA	1482
Db	1441	GAGTTATCTGAAGATATACATAATATTTTATGAAACTATA	1482

RESULT 4	ARI62974	LOCUS	ARI62974	1482 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION	Sequence 45 from patent US 6270771.	ACCESSION	ARI62974				
VERSION	ARI62974.1	GI:	16233440				
KEYWORDS	.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 1482)						
AUTHORS	Guerin-Marchand C. and Druilhé P.						
TITLE	Pepitide sequences specific for the hepatic stages of <i>P. falciparum</i>						
JOURNAL	bearing epitopes capable of stimulating the T lymphocytes						
FEATURES	Patent: US 6270771-A 45 07-AUS-2001;						
source	Location/Qualifiers						
	1..1482						
	/organism="unknown"						
	/mol_type="unassigned DNA"						
ORIGIN							

Query Match	99.9%	Score 1480.4;	DB 6;	Length 1482;
Best Local Similarity	99.9%	Pred. No. 2e-158;		
Matches 1481; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	1	CAGAACAACAAAGCGATCTAGAA	CAAGAGAGACGTGCTAAAGAAAAGTTGCCAAGACA	60
Db	1	CAGAGACAACAAGCGATCTAGAAC	CAAGAGAGACGTGCTAAAGAAAAGTTGCCAAGACA	60
Qy	61	CAAAAGCGATTTAGAACAGATAG	ACTTGGCTTAAAGAAAAGTTTCAAGACAGCAAAAGCGAT	120
Db	61	CAAAAGCGATTTAGAACAGATAG	ACTTGGCTTAAAGAAAAGTTTCAAGACAGCAAAAGCGAT	120
Qy	121	TTAGAACACAGAGAGACTTGCT	TAAAGAAAAGTTGCCAAGAACAAACAAAGCGATCTAGACA	180
Db	121	TTAGAACACAGAGAGACTTGCT	TAAAGAAAAGTTGCCAAGAACAAACAAAGCGATCTAGACA	180
Qy	181	GAGAGACGTGCTTAAAGAAAAG	TTGCCAAGAACAAACACCGATTTTGAACAAGAGAGACGT	240
Db	181	GAGAGACGTGCTTAAAGAAAAG	TTGCCAAGAACAAACACCGATTTTGAACAAGAGAGACGT	240
Qy	241	GCTAAAGAAAAGTTGCTCAAG	AACCAACAAAGCGATTTTGAACAAGATAGACTTGCTTAAAG	300
Db	241	GCTAAAGAAAAGTTGCTCAAG	AACCAACAAAGCGATTTTGAACAAGATAGACTTGCTTAAAG	300
Qy	301	AAATTACAAAGACAGCAAAAG	CGATTTTGAACAAGAGAGACGTGCTTAAAGAAAAGTTGCCA	360
Db	301	AAATTACAAAGACAGCAAAAG	CGATTTTGAACAAGAGAGACGTGCTTAAAGAAAAGTTGCCA	360
Qy	361	GAACAACAAGAGCGATTTTGA	ACAAGAGAGACGTGCTTAAAGAAAAGTTGCCAAGACA	420
Db	361	GAACAACAAGAGCGATTTTGA	ACAAGAGAGACGTGCTTAAAGAAAAGTTGCCAAGACA	420
Qy	421	ACCGATTTTGAACAAGAGAG	ACTTGGCTTAAAGAAAAGTTGCCAAGAACAAACAAAGCGATTTA	480
Db	421	ACCGATTTTGAACAAGAGAG	ACTTGGCTTAAAGAAAAGTTGCCAAGAACAAACAAAGCGATTTA	480

OY	481	GAACAGAGACACGCGCTAAAGAAAAAGTTGCAAGACAAACAACGCAATTTAGAACAGAG	540
Db	481	GAACAGAGACACGCGCTAAAGAAAAAGTTGCAAGACAAACAACGCAATTTAGAACAGAG	540
OY	541	AGACGCGCTAAAGAAAAAGTTGCTCAAGAACAAACAAAGCGATTTTAGAACAGAGACGTCCT	600
Db	541	AGACGCGCTAAAGAAAAAGTTGCTCAAGAACAAACAAAGCGATTTTAGAACAGAGACGTCCT	600
OY	601	AAAGAAAAGTTGCAAGACACGACCAAGAGATTTAGAACAAAGAAAGGCTGATACGAAAAA	660
Db	601	AAAGAAAAGTTGCAAGACACGACCAAGAGATTTAGAACAAAGAAAGGCTGATACGAAAAA	660
OY	661	AAATTTAGAAAAAAGAAAAAGGAAATGAGATATATATAGACAGAGAGATTTATATGTCGTTTA	720
Db	661	AAATTTAGAAAAAAGAAAAAGGAAATGAGATATATATAGACAGAGATTTATATGTCGTTTA	720
OY	721	GAATATACCAAGCTATAGAACTTCCATCAGAAAAATGAACTGTGATATTTATATACCAATCAA	780
Db	721	GAATATACCAAGCTATAGAACTTCCATCAGAAAAATGAACTGTGATATTTATATACCAATCAA	780
OY	781	TCTTCTTTAAGCTCAGGACAAACAGAGGGAATAGTAGAATTTCCAAAGAAATATCTATATA	840
Db	781	TCTTCTTTAAGCTCAGGACAAACAGAGGGAATAGTAGAATTTCCAAAGAAATCTATATATA	840
OY	841	GAAGAAAAAATATAGGAATCTTATTCACAAATGTGTGAAGACGAAGGGAATATCATATAA	900
Db	841	GAAGAAAAAATATAGGAATCTTATTCACAAATGTGTGAAGACGAAGGGAATATCATATAA	900
OY	901	GGACATCTTTGAAGAAAAAGAAAGATGGTTCAATAAAAACGAAACAAAAAGAAATAAATCT	960
Db	901	GGACATCTTTGAAGAAAAAGAAAGATGGTTCAATAAAAACGAAACAAAAAGAAATAAATCT	960
OY	961	GCTGCATATCAAAATCATATCATTTAGAGACAGTAAATATTTCTGATAGTTTAATGATTTTCAA	1020
Db	961	GCTGCATATCAAAATCATATCATTTAGAGACAGTAAATATTTCTGATAGTTTAATGATTTTCAA	1020
OY	1021	ATTAAGTAAGTATGAGAGATGAATAAGTCTGCAATATGACGAATTCATTAATAGATGAGAA	1080
Db	1021	ATTAAGTAAGTATGAGAGATGAATAAGTCTGCAATATGACGAATTCATTAATAGATGAGAA	1080
OY	1081	GAAGATGATGAAGACTTAGACGAATTTAAAGCCTATTTGCAATATGACAAATTTCCAGAT	1140
Db	1081	GAAGATGATGAAGACTTAGACGAATTTAAAGCCTATTTGCAATATGACAAATTTCCAGAT	1140
OY	1141	GAAGAAAAATATGGAATTTATTAAGAACTTAGAAGTTTGATATGAGAAAAATGAAAAATTTA	1200
Db	1141	GAAGAAAAATATGGAATTTATTAAGAACTTAGAAGTTTGATATGAGAAAAATGAAAAATTTA	1200
OY	1201	GATGATTTAGATGAAGAAATAGAAAAATCATCAGAAAGAAATTAATCGAAGAAAAATAAAA	1260
Db	1201	GATGATTTAGATGAAGAAATAGAAAAATCATCAGAAAGAAATTAATCGAAGAAAAATAAAA	1260
OY	1261	AAAGGAAAGAAATATGAAAAAACAAGAGTAATATTTTAAACCAATGATPAAAGTTTG	1320
Db	1261	AAAGGAAAGAAATATGAAAAAACAAGAGTAATATTTTAAACCAATGATPAAAGTTTG	1320
OY	1321	TATGATGAGCAATATTAATAAATAATATAAATGATAAGCAGTTTAATATGAGAAAGGAAAAA	1380
Db	1321	TATGATGAGCAATATTAATAAATAATATAAATGATAAGCAGTTTAATATGAGAAAGGAAAAA	1380
OY	1381	TTTCATPAAATCATTTGTTTCATATATTTTGACGAGACAAATGAAATTTTCAAGTCGTGAT	1440
Db	1381	TTTCATPAAATCATTTGTTTCATATATTTTGACGAGACAAATGAAATTTTCAAGTCGTGAT	1440
OY	1441	GAGTTATCTGAAAGATATACTAAATATTTTATGAAACTATATA	1482
Db	1441	GAGTTATCTGAAAGATATACTAAATATTTTATGAAACTATATA	1482

[illegible]

ACCESSION	ARI178263
VERSION	ARI178263.1
KEYWORDS	GI:20219401
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 1482)
TITLE	Guerin-Marchand,C. and Druilhe,P. Peptide sequences specific for the hepatic stages of <i>P. falciparum</i> bearing epitopes capable of stimulating the T lymphocytes
JOURNAL	Patent: US 6319502-A 45 20-NOV-2001;
FEATURES	Location/Qualifiers
source	1..1482 /organism="Unknown" /mol_type="unassigned DNA"

Query Match	99.9%	Score 1480.4	DB 6	Length 1482
Best Local Similarity	99.9%	Pred. No. 2e+158		
Matches 1481	0	Mismatches 1	Indels 0	Gaps 0

OY	1	CAAGACAACAAGGATCTTGAACAACAAGAGAGAGCTGCTAAAGAAAAGTTGCCAAGACA	60
Db	1	CAGACACAACAAGCGATCTTGAACACAGAGAGAGCTGCTAAAGAAAAGTTGCCAAGACA	60
OY	61	CAAAAGCATTTAGAACAAAGATAGACTTGTCTAAAGAAAAGTTACAGAGACAGCAAGCGAT	120
Db	61	CAAAAGCATTTAGAACAAAGATAGACTTGTCTAAAGAAAAGTTACAGAGACAGCAAAAGCGAT	120
OY	121	TTTAGAACACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAGAACAAAGAGGATCTGAGACAA	180
Db	121	TTTAGAACACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAGAACAAAGAGGATCTGAGACAA	180
OY	181	GAGAGACGTGCTAAAGAAAAGTTGCAGAGAACAAACAAAGCGATTTTAGAACACAGAGAGAGCT	240
Db	181	GAGAGACGTGCTAAAGAAAAGTTGCAGAGAACAAACAAAGCGATTTTAGAACACAGAGAGAGCT	240
OY	241	GCTAAAGAAAAGTTGCAGAGAACAAACAAAGCGATTTAGAACAGATAGACTTGTCTAAAGAA	300
Db	241	GCTAAAGAAAAGTTGCAGAGAACAAACAAAGCGATTTAGAACAGATAGACTTGTCTAAAGAA	300
OY	301	AAAGTTACAAAGACAGCAAAAGCGATTTAGAACACAGAGAGAGCTGCTAAAGAAAAGTTGC	360
Db	301	AAAGTTACAAAGACAGCAAAAGCGATTTAGAACACAGAGAGAGCTGCTAAAGAAAAGTTGC	360
OY	361	GAAACAACAAGGATTTAGAACACAGAGAGAGAGCTGCTAAAGAAAAGTTGCCAAGAACAA	420
Db	361	GAAACAACAAGGATTTAGAACACAGAGAGAGAGCTGCTAAAGAAAAGTTGCCAAGAACAA	420
OY	421	AGCGATTTAGAACACAGAGAGAGCTTGTCTAAAGAAAAGTTGCAGAGAACAAACGCGATTTA	480
Db	421	AGCGATTTAGAACACAGAGAGAGCTTGTCTAAAGAAAAGTTGCAGAGAACAAACGCGATTTA	480
OY	481	GAAACAAGAGAGCGTCTTAAAGAAAAGTTGCAGAGAACAAAGCGATTTTGAACAAGAG	540
Db	481	GAAACAAGAGAGCGTCTTAAAGAAAAGTTGCAGAGAACAAAGCGATTTTGAACAAGAG	540
OY	541	AGAGCGTCTTAAAGAAAAGTTGCAGAGAACAAACAAAGCGATTTAGAACACAGAGAGAGCTGCT	600
Db	541	AGAGCGTCTTAAAGAAAAGTTGCAGAGAACAAACAAAGCGATTTTGAACAAGAGAGAGCTGCT	600
OY	601	AAAGAAAAGTTGCAGAGAGCAAGCAAGATTTAGAACAAAGGAAAGGCTGATCGAAAAAA	660
Db	601	AAAGAAAAGTTGCAGAGAGCAAGCAAGATTTAGAACAAAGGAAAGGCTGATCGAAAAAA	660
OY	661	AATTTAGAAAAGAAAAGGAAACATGAGATATTTAGACAGAGATTTATATGCTGCTTAA	720
Db	661	AATTTAGAAAAGAAAAGGAAACATGAGATATTTAGACAGAGATTTATATGCTGCTTAA	720
OY	721	GAAATACAGACTATGAACTTCCATCAGAAAAATGAACTGATATTTATATACACATCAA	780
Db	721	GAAATACAGACTATGAACTTCCATCAGAAAAATGAACTGATATTTATATACACATCAA	780

QY 781 TCTTCTTACCTCAGACACAGAGGAAATAGATAGATTCAGAGAAATATCTATAA 840
DB 781 TCTTCTTACCTCAGACACAGAGGAAATAGATAGATTCAGAGAAATATCTATAA 840
QY 841 GAAAAACAATAAGAAATCTATTACAACTGTTGAAGAGAGGAGATATACATAA 900
DB 841 GAAAAACAATAAGAAATCTATTACAACTGTTGAAGAGAGGAGATATACATAA 900
QY 901 GGACATCTTGAAGAAAAAGAGATGTTCAATAAAAACCAAGCAAAAAAGATAACT 960
DB 901 GGACATCTTGAAGAAAAAGAGATGTTCAATAAAAACCAAGCAAAAAAGATAACT 960
QY 961 GCTGACATACAAATCATCATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTCA 1020
DB 961 GCTGACATACAAATCATCATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTCA 1020
QY 1021 ATTAAGTAAATAGAGATGAAATTAAGTCTGAATTAAGAGATTCATTAATAGTAAGA 1080
DB 1021 ATTAAGTAAATAGAGATGAAATTAAGTCTGAATTAAGAGATTCATTAATAGTAAGA 1080
QY 1081 GAAAGTGAAGAACTTAGACGAATTTAAGCCTATTGTCATATAGCAATTTCCAGAT 1140
DB 1081 GAAAGTGAAGAACTTAGACGAATTTAAGCCTATTGTCATATAGCAATTTCCAGAT 1140
QY 1141 GAAAGAAACATAGAAATTTAATAAGAACTAGAAAGATTTGATAGAGAAAAATGAATTTA 1200
DB 1141 GAAAGAAACATAGAAATTTAATAAGAACTAGAAAGATTTGATAGAGAAAAATGAATTTA 1200
QY 1201 GATGATTTAGATGAAGAAATAGAAAAATCATCAGAAATTAATCTGAAGAAAAATTA 1260
DB 1201 GATGATTTAGATGAAGAAATAGAAAAATCATCAGAAATTAATCTGAAGAAAAATTA 1260
QY 1261 AAAAGAAAGAAATAGAAAAACAAAGGATTAATTTTAAACCAATGATTAAGCTTTG 1320
DB 1261 AAAAGAAAGAAATAGAAAAACAAAGGATTAATTTTAAACCAATGATTAAGCTTTG 1320
QY 1321 TATGATGAGCATATTAATAAATATTAATAAGTAAAGCAGTTAATAGAGAAAAA 1380
DB 1321 TATGATGAGCATATTAATAAATATTAATAAGTAAAGCAGTTAATAGAGAAAAA 1380
QY 1381 TTCTATTAATCATTTGTTTCAATATTTGAAGAGCAATGAATTTTACAGATCGTGAT 1440
DB 1381 TTCTATTAATCATTTGTTTCAATATTTGAAGAGCAATGAATTTTACAGATCGTGAT 1440
QY 1441 GAGTTATCTGAAGATTAATTAATTTTATGAACCTATA 1482
DB 1441 GAGTTATCTGAAGATTAATTAATTTTATGAACCTATA 1482
RESULT 6
A28743 1493 bp DNA linear PAT 04-JUN-1995
LOCUS A28743
DEFINITION CDNA for USA gene 3' (partial).
ACCESSION A28743
VERSION A28743.1 GI:1247516
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1493)
AUTHORS
TITLE LIVER-STAGE-SPECIFIC PEPTIDE SEQUENCES OF P. FALCIPARUM BEARING
JOURNAL EPITOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES
FEATURES Patent: WO 9213884-A 32 20-AUG-1992;
source Location/Qualifiers
1..1493
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 97.6%; Score 1446; DB 6; Length 1493;
Best Local Similarity 99.8%; Pred. No. 1.5e-154;
Matches 1479; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
QY 1 CAAGAACAAACAAGCGATCTTGAAACAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAAC 60
DB 1 CAAGAACAAACAAGCGATCTTGAAACAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAAC 60
QY 61 CAAGCCGATTTTGAACAAAGATAGACTTGCTTAAAGAAAGTTTACAGAGCAGCAAGCGAT 120
DB 61 CAAGCCGATTTTGAACAAAGATAGACTTGCTTAAAGAAAGTTTACAGAGCAGCAAGCGAT 120
QY 121 TTAGAACAAAGAGAGCTTGCTTAAAGAAAGTTTGAACAAACAAGCGATCTTGAACAA 180
DB 121 TTAGAACAAAGAGAGCTTGCTTAAAGAAAGTTTGAACAAACAAGCGATCTTGAACAA 179
QY 181 GAGAGACGCTGCTTAAAGAAAGTTGCAAGAACAAACAAGCGATTTTGAACAAAGAGAGCT 240
DB 181 GAGAGACGCTGCTTAAAGAAAGTTGCAAGAACAAACAAGCGATTTTGAACAAAGAGAGCT 239
QY 241 GCTAAAGAAAGTTGCAAGAACAAACAAGCGATTTTGAACAAAGAGAGCTTGAACAA 300
DB 241 GCTAAAGAAAGTTGCAAGAACAAACAAGCGATTTTGAACAAAGAGAGCTTGAACAA 299
QY 301 AAGTTACAAAGAGCAGCAAGCGATTTTGAACAAAGAGAGAGCTTGAACAAAGAGTTGCA 360
DB 301 AAGTTACAAAGAGCAGCAAGCGATTTTGAACAAAGAGAGAGCTTGAACAAAGAGTTGCA 359
QY 361 GAAACAACAAGCGATTTTGAACAAAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAA 420
DB 361 GAAACAACAAGCGATTTTGAACAAAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAA 418
QY 421 AGCGATTTTGAACAAAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAAAGCGATTTA 480
DB 421 AGCGATTTTGAACAAAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAAAGCGATTTA 478
QY 479 GAAACAAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAAAGCGATTTTGAACAAAG 538
DB 479 GAAACAAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAAAGCGATTTTGAACAAAG 538
QY 541 AGAGCTGCTTAAAGAAAGTTGCAAGAACAAACAAGCGATTTTGAACAAAGAGAGAGCTGCT 600
DB 541 AGAGCTGCTTAAAGAAAGTTGCAAGAACAAACAAGCGATTTTGAACAAAGAGAGAGCTGCT 597
QY 597 AGAGCTGCTTAAAGAAAGTTGCAAGAACAAACAAGCGATTTTGAACAAAGAGAGAGCTGCT 597
DB 597 AGAGCTGCTTAAAGAAAGTTGCAAGAACAAACAAGCGATTTTGAACAAAGAGAGAGCTGCT 597
QY 601 AAAAGAAAGTTGCAAGAGAGAGAGATTTTGAACAAAGAGAGAGCTGATTCAGAAAAA 660
DB 601 AAAAGAAAGTTGCAAGAGAGAGAGATTTTGAACAAAGAGAGAGCTGATTCAGAAAAA 657
QY 657 AAAAGAAAGTTGCAAGAGAGAGAGATTTTGAACAAAGAGAGAGCTGATTCAGAAAAA 657
DB 657 AAAAGAAAGTTGCAAGAGAGAGAGATTTTGAACAAAGAGAGAGCTGATTCAGAAAAA 657
QY 721 GAAATACCAAGCTATAGAACTTCCATCAGAAAAATGAACCTGATATTTATACCATCA 780
DB 721 GAAATACCAAGCTATAGAACTTCCATCAGAAAAATGAACCTGATATTTATACCATCA 777
QY 777 GAAATACCAAGCTATAGAACTTCCATCAGAAAAATGAACCTGATATTTATACCATCA 777
DB 777 GAAATACCAAGCTATAGAACTTCCATCAGAAAAATGAACCTGATATTTATACCATCA 777
QY 781 TCTTCTTACCTCAGAGACACAGAGGAAATAGATAGATTCAGAGAAATATCTATAA 840
DB 781 TCTTCTTACCTCAGAGACACAGAGGAAATAGATAGATTCAGAGAAATATCTATAA 837
QY 841 GAAAAACAATAAGAAATCTATTACAACTGTTGAAGAGAGGAGATATACATAA 900
DB 841 GAAAAACAATAAGAAATCTATTACAACTGTTGAAGAGAGGAGATATACATAA 897
QY 901 GGACATCTTGAAGAAAAAGAGATGTTCAATAAAAACCAAGCAAAAAAGATAACT 960
DB 901 GGACATCTTGAAGAAAAAGAGATGTTCAATAAAAACCAAGCAAAAAAGATAACT 957
QY 961 GCTGACATACAAATCATCATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTCA 1020
DB 961 GCTGACATACAAATCATCATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTCA 1017
QY 1021 ATTAAGTAAATAGAGATGAAATTAAGTCTGAATTAAGAGATTCATTAATAGTAAGA 1080

|||||
Db 1018 ATAAAGTAAGTATGAGATGAAATTAAGTGTGAATATGACATTCATTAAATAGATGAAGAA 1077
|||
Qy 1081 GAAGATGATGAAGACTTAGACGAATTTTACCTATTGTGCAATATGACAAATTTCCAGAT 1140
|||
Db 1078 GAAGATGATGAAGACTTAGACGAATTTTACCTATTGTGCAATATGACAAATTTCCAGAT 1137
|||
Qy 1141 GAAGAAAACATAGGAATTTTAAAGAACTAGAGATTTGATAGAGAAAATGAAAATTTTA 1200
|||
Db 1138 GAAGAAAACATAGGAATTTTAAAGAACTAGAGATTTGATAGAGAAAATGAAAATTTTA 1197
|||
Qy 1201 GATGATTTAGATGAGGAATTAAGAAAATCATCGAAGAAATTAATCTGAGAAAATTAATA 1260
|||
Db 1198 GATGATTTAGATGAGGAATTAAGAAAATCATCGAAGAAATTAATCTGAGAAAATTAATA 1257
|||
Qy 1261 AAAGGAAAGAAATATGAAAAACAAAGATTAATTTTAAACCAATGATTAAGAAAGTTTG 1320
|||
Db 1258 AAAGGAAAGAAATATGAAAAACAAAGATTAATTTTAAACCAATGATTAAGAAAGTTTG 1317
|||
Qy 1321 TATGATGAGCATATTAATAAATATATAATGATTAAGAGGTTAATTAAGAAAAGAAAA 1380
|||
Db 1318 TATGATGAGCATATTAATAAATATATAATGATTAAGAGGTTAATTAAGAAAAGAAAA 1377
|||
Qy 1381 TTCTATAAATCATTTGTTTCTATATTTTGACGAGACAAATGAAATTTTACAGATCGTGGAT 1440
|||
Db 1378 TTCTATAAATCATTTGTTTCTATATTTTGACGAGACAAATGAAATTTTACAGATCGTGGAT 1437
|||
Qy 1441 GAGTTATCTGAGATATTAATTAATTTTATGAAACTATTA 1482
|||
Db 1438 GAGTTATCTGAGATATTAATTAATTTTATGAAACTATTA 1479
|||

RESULT 7
ARI62967 1493 bp DNA linear PAT 17-OCT-2001
LOCUS Sequence 38 from patent US 6270771.
DEFINITION ARI62967
ACCESSION ARI62967
VERSION ARI62967.1 GI:16233432
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1493)
AUTHORS
TITLE
JOURNAL
FEATURES
Source
1. .1493
Location/Qualifiers
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 97.6%; Score 1446; DB 6; Length 1493;
Beet Local Similarity 99.8%; Pred. No. 1.5e-154;
Matches 1479; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

ORIGIN
Qy 1 CAAAGCAACAAGGATCTAGAACAAAGAGACGTGCTAAAGAAAAGTTGCAAGAACAA 60
|||
Db 1 CAAAGCAACAAGGATCTAGAACAAAGAGACGTGCTAAAGAAAAGTTGCAAGAACAA 60
|||
Qy 61 CAAAGCAATTTAGAACAAAGATAGACTTGTCTAAAGAAAAGTTTACAAGACCAAGCGAT 120
|||
Db 61 CAAAGCAATTTAGAACAAAGATAGACTTGTCTAAAGAAAAGTTTACAAGACCAAGCGAT 120
|||
Qy 121 TTAGAACAAAGAGACTTGTCTAAAGAAAAGTTGCAAGAACAAAGGATCTTGAACAA 180
|||
Db 121 TTAGAACAAAGAGACTTGTCT - AAGAAAAGTTGCAAGAACAAAGGATCTTGAACAA 179
|||
Qy 181 GAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAAGGATTTTGAACAAAGAGACGT 240
|||
Db 180 GAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAAGGATTTTGAACAAAGAGACGT 239
|||

Qy 241 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAGATAGACTTGTCTAAAGAA 300
|||
Db 240 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAGATAGACTTGTCTAAAGAA 299
|||
Qy 301 AAGTTTCAAGAGACGCAAGGCGATTTTGAACAAGAGAGACGTGCTTAAAGAAAAGTTGCAA 360
|||
Db 300 AAGTTTCAAGAGACGCAAGGCGATTTTGAACAAGAGAGACGTGCTTAAAGAAAAGTTGCAA 359
|||
Qy 361 GAACAAACAAGCGATTTTGAACAAGAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAA 420
|||
Db 360 GAACAAACAAGCGATTTTGAACAAGAGAGACGTGCT - AAGAAAAGTTGCAAGAACAA 418
|||
Qy 421 AGCGATTTTGAACAAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTA 480
|||
Db 419 AGCGATTTTGAACAAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTA 478
|||
Qy 481 GAACAGAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAGAG 540
|||
Db 479 GAACAGAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAGAG 538
|||
Qy 541 AGACGTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAGAGAGACGTGCT 600
|||
Db 539 AGACGTGCT - AAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAGAGAGACGTGCT 597
|||
Qy 601 AAAGAAAAGTTGCAAGAGACGCAAGGATTTTGAACAAGGAAAGGCTGATACGAAAAAA 660
|||
Db 598 AAAGAAAAGTTGCAAGAGACGCAAGGATTTTGAACAAGGAAAGGCTGATACGAAAAAA 657
|||
Qy 661 AATTTGAAAAGAAAAGGAAACATGAGATATTTTGAAGAGGATTTTATATGCTCGTTTA 720
|||
Db 658 AATTTGAAAAGAAAAGGAAACATGAGATATTTTGAAGAGGATTTTATATGCTCGTTTA 717
|||
Qy 721 GAAATACACGCTATTAAGACTTCCATCAAGAAATGAAACGTGATATTTATACACATCAA 780
|||
Db 718 GAAATACACGCTATTAAGACTTCCATCAAGAAATGAAACGTGATATTTATACACATCAA 777
|||
Qy 781 TCTTCTTTACCTCAGACCAACAGAGGAAATGATGAGATTTCCAAAGGAAATATCTATATA 840
|||
Db 778 TCTTCTTTACCTCAGACCAACAGAGGAAATGATGAGATTTCCAAAGGAAATATCTATATA 837
|||
Qy 841 GAAAAAACAAATAGAGATCTATTTACAAACAATGTTGAAGACGAAGGATATACATTA 900
|||
Db 838 GAAAAAACAAATAGAGATCTATTTACAAACAATGTTGAAGACGAAGGATATACATTA 897
|||
Qy 901 GACATCTTGAAGAAAAGAAAGTGTGTTCAATTAACACGAACAAAGAAAGATTAATCT 960
|||
Db 898 GACATCTTGAAGAAAAGAAAGTGTGTTCAATTAACACGAACAAAGAAAGATTAATCT 957
|||
Qy 961 GCTGACATTAACAATCATTAATTAAGACAGTAAATATTTCTGATGTTAATGATTTTCAA 1020
|||
Db 958 GCTGACATTAACAATCATTAATTAAGACAGTAAATATTTCTGATGTTAATGATTTTCAA 1017
|||
Qy 1021 ATAAAGTAAGTATGAGATGAATTAAGTCTGAATATGACGATTCATTAATATGATGAAGA 1080
|||
Db 1018 ATAAAGTAAGTATGAGATGAATTAAGTCTGAATATGACGATTCATTAATATGATGAAGA 1077
|||
Qy 1081 GAAGATGATGAAGACTTAGACGAATTTTACCTATTGTGCAATATGACAAATTTCCAGAT 1140
|||
Db 1078 GAAGATGATGAAGACTTAGACGAATTTTACCTATTGTGCAATATGACAAATTTCCAGAT 1137
|||
Qy 1141 GAAGAAAACATAGGAATTTTAAAGAACTAGAGATTTGATAGAGAAAATGAAAATTTTA 1200
|||
Db 1138 GAAGAAAACATAGGAATTTTAAAGAACTAGAGATTTGATAGAGAAAATGAAAATTTTA 1197
|||
Qy 1201 GATGATTTAGATGAGGAATTAAGAAAATCATCGAAGAAATTAATCTGAGAAAATTAATA 1260
|||
Db 1198 GATGATTTAGATGAGGAATTAAGAAAATCATCGAAGAAATTAATCTGAGAAAATTAATA 1257
|||
Qy 1261 AAAGGAAAGAAATATGAAAAACAAAGATTAATTTTAAACCAATGATTAAGAAAGTTTG 1320
|||
Db 1258 AAAGGAAAGAAATATGAAAAACAAAGATTAATTTTAAACCAATGATTAAGAAAGTTTG 1317
|||
Qy 1321 TATGATGAGCATATTAATAAATATATAATGATTAAGAGGTTAATTAAGAAAAGAAAA 1380
|||

|||||
Db 1318 TATGATGAGCATATTTAAAAATATTAAGATTAAGCAGCTTAATATAGGAAAAAGAAAA 1377
|||
Qy 1381 TTCATTAATATCTTTGTTTATATATTTTGACGAGACATGAATTTTACAGATCGTGAT 1440
|||
Db 1378 TTCATTAATATCTTTGTTTATATATTTTGACGAGACATGAATTTTACAGATCGTGAT 1437
|||
Qy 1441 GAGTTATCTGAAGATTAATCTAAATTTTATGAACCTATA 1482
|||
Db 1438 GAGTTATCTGAAGATTAATCTAAATTTTATGAACCTATA 1479
|||
RESULT 8
AR178256 1493 bp DNA linear PAT 20-APR-2002
LOCUS Sequence 38 from patent US 6319502.
DEFINITION AR178256
VERSION AR178256.1 GI:20219394
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1493)
AUTHORS Guerin-Marchand,C. and Druilhe,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
JOURNAL bearing epitopes capable of stimulating the T lymphocytes
FEATURES
Location/Qualifiers
1..1493
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 97.6%; Score 1446; DB 6; Length 1493;
Best Local Similarity 99.8%; Pred.No.1.5e-154;
Matches 1479; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
1 CAAGAACACAAAGCATCTAGAACAAAGACGCTAAAGAAAAAGTTGCAGAACAA 60
1 CANGAACACAAAGCATCTAGAACAAAGACGCTAAAGAAAAAGTTGCAGAACAA 60
Qy 61 CAAGCGATTTAGAACAAAGATAGACTTGTAAAGAAAGTTACAGAGCAGCAAGCGAT 120
Db 61 CAAGCGATTTAGAACAAAGATAGACTTGTAAAGAAAGTTACAGAGCAGCAAGCGAT 120
Qy 121 TTAGAACAAAGAGACTTGTCTTAAGAAAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
Db 121 TTAGAACAAAGAGACTTGTCT- AAGAAAAAGTTGCAAGAACAAAGCGATCTAGAACAA 179
Qy 181 GAGAGCGTGTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCT 240
Db 180 GAGAGCGTGTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCT 239
Qy 241 GCTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGTAAAGAA 300
Db 240 GCTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGTAAAGAA 299
Qy 301 AAGTTCAAGAGAGCAAGCGATTTAGAACAAAGAGAGAGCTGTAAAGAAAAAGTTGCAA 360
Db 300 AAGTTCAAGAGAGCAAGCGATTTAGAACAAAGAGAGAGCTGTAAAGAAAAAGTTGCAA 359
Qy 361 GAAACAACAAGCGATTTAGAACAAAGAGAGAGCTGTAAAGAAAAAGTTGCAAGAACAA 420
Db 360 GAAACAACAAGCGATTTAGAACAAAGAGAGAGCTGT- AAGAAAAAGTTGCAAGAACAA 418
Qy 421 AGCGATTTAGAACAAAGAGAGACTTGTAAAGAAAAAGTTGCAAGAACAAAGCGATTTA 480
Db 419 AGCGATTTAGAACAAAGAGAGACTTGTAAAGAAAAAGTTGCAAGAACAAAGCGATTTA 478
Qy 481 GAAACAAGAGAGAGCTGTCTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAG 540
Db 479 GAAACAAGAGAGAGCTGTCTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAG 538

Qy 541 AGACGCTCTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCT 600
Db 539 AGACGCTCT- AAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCT 597
Qy 601 AAAGAAAAAGTTGCAAGAGCAGCAGCAAGAGATTTAGAACAAAGAAAGCTGATTCGAAAAA 660
Db 598 AAAGAAAAAGTTGCAAGAGCAGCAGCAAGAGATTTAGAACAAAGAAAGCTGATTCGAAAAA 657
Qy 661 AATTAGAAAAAGAAAAAGAAATGAGAGATTAATTTAGCAGAGATTTTATATGTCGTTTA 720
Db 658 AATTAGAAAAAGAAAAAGAAATGAGAGATTAATTTAGCAGAGATTTTATATGTCGTTTA 717
Qy 721 GAAATACCAAGCTATAGAACCTTCATCAGAAAAAGACGAGATTAATATATACATCA 780
Db 718 GAAATACCAAGCTATAGAACCTTCATCAGAAAAAGACGAGATTAATATATACATCA 777
Qy 781 TCTTTTCTTACCTCAGGACAAACAGAGGAAATGATAGAGATTTCCAGAAAAATCTTAATA 840
Db 778 TCTTTCTTACCTCAGGACAAACAGAGGAAATGATAGAGATTTCCAGAAAAATCTTAATA 837
Qy 841 GAAAAACAAATPAGGAATCTATTACAACAATGTTGAAGGACGAAGGATATACATATA 900
Db 838 GAAAAACAAATPAGGAATCTATTACAACAATGTTGAAGGACGAAGGATATACATATA 897
Qy 901 GCACATCTTGAAGAAAAAGATGCTTCAATTAACCAAGAACAAAGAAAGATTAATCT 960
Db 898 GCACATCTTGAAGAAAAAGATGCTTCAATTAACCAAGAACAAAGAAAGATTAATCT 957
Qy 961 GCTGACATACAAATCATACATTTAGAGACAGTAAATATTTTCGATGTTTAATGATTTTCAA 1020
Db 958 GCTGACATACAAATCATACATTTAGAGACAGTAAATATTTTCGATGTTTAATGATTTTCAA 1017
Qy 1021 ATTAAGTATGATGAGAGATGAATAAGCTGTAATGACGATTCATTAATAGATGAAGA 1080
Db 1018 ATTAAGTATGATGAGAGATGAATAAGCTGTAATGACGATTCATTAATAGATGAAGA 1077
Qy 1081 GAAGATGATGAAGACTTACAGCAATTTAAGCTATTTGCAATATGACAAATTTCCAAGAT 1140
Db 1078 GAAGATGATGAAGACTTACAGCAATTTAAGCTATTTGCAATATGACAAATTTCCAAGAT 1137
Qy 1141 GAAAGAAACATAGGAATTTATTAAGAACTAGAGAAATTTGATAGAAAAATGAATAATTTA 1200
Db 1138 GAAAGAAACATAGGAATTTATTAAGAACTAGAGAAATTTGATAGAAAAATGAATAATTTA 1197
Qy 1201 GATGATTTAGATGAAGAAATGAAGAAATCATCAGAGAAATTTCTGAAGAAAAATTAATA 1260
Db 1198 GATGATTTAGATGAAGAAATGAAGAAATCATCAGAGAAATTTCTGAAGAAAAATTAATA 1257
Qy 1261 AAAGAAAAATATGAAAAAAACAAAGATTAATTTTAAACCAATGATTAAGATTTTG 1320
Db 1258 AAAGAAAAATATGAAAAAAACAAAGATTAATTTTAAACCAATGATTAAGATTTTG 1317
Qy 1321 TATGATGAGCATTTTAAAAATATTAATAATGTAAGCAGGTTAAATAGAAAAAGAAAA 1380
Db 1318 TATGATGAGCATTTTAAAAATATTAATAATGTAAGCAGGTTAAATAGAAAAAGAAAA 1377
Qy 1381 TTCATTAATATCTTTGTTTATATATTTTGACGAGACAAATGAATTTTACAGATCGTGAT 1440
Db 1378 TTCATTAATATCTTTGTTTATATATTTTGACGAGACAAATGAATTTTACAGATCGTGAT 1437
Qy 1441 GAGTTATCTGAAGATTAATCTAAATTTTATGAACCTATA 1482
Db 1438 GAGTTATCTGAAGATTAATCTAAATTTTATGAACCTATA 1479
|||
RESULT 9
PFLSAIG 5970 bp DNA linear INV 18-APR-2005
LOCUS P.falciparum LSA-1 gene for liver stage antigen.
DEFINITION X56203 S75010 S75012 S75014 S75016
ACCESSION X56203.1 GI:9915
VERSION
KEYWORDS liver stage antigen.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)

REFERENCE	ORGANISM
1	Plasmodium falciparum
2	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
3	Zhu, J. and Hollingdale, M. R.
4	Structure of Plasmodium falciparum liver stage antigen-1
5	Mol. Biochem. Parasitol. 46 (2), 223-226 (1991)
6	1840628
7	2 (bases 1 to 5970)
8	Zhu, J. and Hollingdale, M. R.
9	Direct Submission
10	Submitted (28-NOV-1990) Zhu, J., Hollingdale M. Biomedical Research
11	Institute, 1211 Parklawn Drive, Rockville, MD 20852, U.S.A
12	On or before Jun 27, 2005 this sequence version replaced gi:241572,
13	gi:241573, gi:241574, gi:241575.
14	Location/Qualifiers
15	1..5970
16	/organism="Plasmodium falciparum"
17	/mol_type="Genomic DNA"
18	/strain="NF54"
19	/db_xref="taxon:5833"
20	/clone="RI-LISA segment, DRAI-LISA segment"
21	/cell_line="NF54 blood stage culture"
22	/clone_1lb="pUC13 plasmid, RI-sublib, DRAI-sublib"
23	/dev_stage="liver stage"
24	30..36
25	79..5808
26	/gene="LSA-1"
27	79..5808
28	/gene="LSA-1"
29	/codon_start=1
30	/product="liver stage antigen"
31	/protein_id="CAA39663.1"
32	/db_xref="GI:3916"
33	/db_xref="HSP:1D7M"
34	/db_xref="UniProt/TREMBL:Q25893"
35	/translation="MKHLIYSPFLIVNLILPIHNGKIIINSKDRILITSNRSQSS
36	NSNRINDEKHEKQVLSHNSVEYTKNNENKPFDDKDELMSVKNVSVQTFSLRL
37	NLGVSENLFLKENLKNGKLIIEHI INDDDKKKYIGQDENRQDELKKAAKTLQGLQ
38	QOSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQ
39	ERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
40	LOEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSD
41	LEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
42	KEKTLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSD
43	QSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
44	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
45	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
46	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
47	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
48	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
49	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
50	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
51	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
52	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
53	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
54	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
55	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
56	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
57	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
58	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
59	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
60	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
61	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
62	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
63	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
64	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
65	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
66	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
67	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
68	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
69	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
70	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
71	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
72	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
73	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
74	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
75	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
76	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
77	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
78	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
79	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
80	EOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQEOQSDLEQERLAKETLQ
81	EOQSDLEQERLAKETLQEOQSDLEQERL

	Matches	145/;	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
Qy	1	CAAGAACAA	CAAAAGCGATCTTAGAA	CAAGAGAGACGTGCTTAAAGAAAAGTTGCAGAA	CAA	60				
Db	4327	CAAGAACAA	CAAAAGCGATCTTAGAA	CAAGAGAGACGTGCTTAAAGAAAAGTTGCAGAA	CAA	4386				
Qy	61	CAAAAGCATTTAGAA	CAAGATAGACTCTTGAAGAAAAGTTTACAGAGACGAAAGCGAT	120						
Db	4387	CAAAAGCATTTAGAA	CAAGAGAGAGACGTGCTTAAAGAAAAGTTTGCAGAGAACAA	CAAAAGCGAT	4446					
Qy	121	TTAGAACAA	GAGAGACTTGTCTTAAAGAAAAGTTGCAGAA	CAACAAAGCGATCTTAGAA	CAA	180				
Db	4447	TTTAGAACAA	GAGAGACGTGCTTAAAGAAAAGTTGCAGAA	CAACAAAGCGATTTTAGAACAA	4506					
Qy	181	GAGAGACGTGCTTAAAGAAAAGTTGCAGAA	CAACAAAGCGATTTAGAACAGAGACGT	240						
Db	4507	GAGAGACTTGTCTTAAAGAAAAGTTGCAGAA	CAACAAAGATTTAGAACAGAGAGACGT	4566						
Qy	241	GCTTAAAGAAAAGTTGCAGAA	CAACAAAGGATTTTGAACAGATGACTTGTCTTAAAGAA	300						
Db	4567	GCTTAAAGAAAAGTTGCAGAA	CAACAAAGGATTTTGAACAGAGACAGACGTGCTTAAAGAA	4626						
Qy	301	AAGTTAACAA	GAGACGCAACGATTTTGAACAGAGAGACGTGCTTAAAGAAAAGTTGCAA	360						
Db	4627	AAGTTGCAGAA	CAACAAAGCGATTTAGAACAGAGAGACCTTGCTAATGAAAAGTTGCAA	4686						
Qy	361	GAACAA	CAAAAGCGATTTTGAACAGAGAGACGTGCTTAAAGAAAAGTTGCAGAA	CAA	420					
Db	4687	GAGCAACAA	AGAGATTTAGAACAAAGAGACGTGCTTAAAGAAAAGTTGCAGAA	CAACAA	4746					
Qy	421	AGCGATTTAGAA	CAAGAGACCTTGTCTTAAAGAAAAGTTGCAGAA	CAACAAAGCGATTTA	480					
Db	4747	AGCGATTTTAA	CAACAAAGAGACCGTCTTAAAGAAAAGTTGCAGAA	CAACAAAGCGATTTA	4806					
Qy	481	GAACAA	GAGAGACGTGCTTAAAGAAAAGTTGCAGAA	CAACAAAGCGATTTTGAACAA	540					
Db	4807	GAACAA	GAGAGACGTGCTTAAAGAAAAGTTGCAGAA	CAACAAAGCGATTTTGAACAA	4866					
Qy	541	AGACGTGCTTAAAGAAAAGTTGCAGAA	CAACAAAGCGATTTTGAACAGAGAGACGTGCT	600						
Db	4867	AGACTTGTCTTAAAGAAAAGTTGCAGAA	CAACAAAGATTTTGAACAGAGAGACGTGCT	4926						
Qy	601	AAAGAAAAGTTGCAGAGACGCAACAA	AGATTTTGAACAAAGGAAGCTGATTCGAAA	660						
Db	4927	AAAGAAAAGTTGCAGAGACGCAACAA	AGATTTTGAACAAAGGAAGCTGATTCGAAA	4986						
Qy	661	AATTTAGAAA	AAAAAGGAACATGAGATATATTTACAGAGATTTATATATGATGCTTTA	720						
Db	4987	AATTTAGAAA	AAAAAGGAACATGAGATATATTTACAGAGATTTATATATGATGCTTTA	5046						
Qy	721	GAATATAC	AGATATAGAACTTTCATTCAGAAAATGAA	CGTGGATTTATATACCAATCAA	780					
Db	5047	GAATATAC	AGATATAGAACTTTCATTCATTCAGAAAATGAA	CGTGGATTTATATACCAATCAA	5106					
Qy	781	TCTTCTTT	ACTTCAGAGCAACAGAGGAATATATAGATTTCAAGGAATATCTATATATA	840						
Db	5107	TCTTCTTT	ACTTCAGAGCAACAGAGGAATATATAGATTTCAAGGAATATCTATATATA	5166						
Qy	841	GAAGAAA	CAAAATGAGAACTTATTA	CAACAAATGTTAAGAGACGAAGGATATCATATA	900					
Db	5167	GAAGAAA	CAAAATGAGAACTTATTA	CAACAAATGTTAAGAGACGAAGGATATCATATA	5226					
Qy	901	GGAATCTT	GAGAAAAGATGTTTCAATATA	CAAGAACAAAGAAAGATATAATCT	960					
Db	5227	GGAATCTT	GAGAAAAGATGTTTCAATATA	CAAGAACAAAGAAAGATATAATCT	5286					
Qy	961	GCTGACAT	CAACAAATCATATTA	GAGACAGTAAATATTTCTGATGTTAATGATTTTCAA	1020					
Db	5287	GCTGACAT	CAACAAATCATATTA	GAGACAGTAAATATTTCTGATGTTAATGATTTTCAA	5346					
Qy	1021	ATAAGTA	AGTATGAGATGAAATTAAGTGCTGAATATATACATTCATTAATATAGATGAAGAA	1080						
Db	5347	ATAAGTA	AGTATGAGATGAAATTAAGTGCTGAATATATACATTCATTAATATAGATGAAGAA	5406						

QY	1081	GAAGATGATGAAGACTTGAAGCAATTTAAAGCCTATTTGCAATATGACAAATTTCCAAAGAT	1140
Db	5407	GAAGATGATGAAGACTTGAAGCAATTTAAAGCCTATTTGCAATATGACAAATTTCCAAAGAT	5466
QY	1141	GAAGAAAACATAGGAATTTATTAAGAACTAGCAAGTTTGAATAGGAAAAATGAATAATTTA	1200
Db	5467	GAAGAAAACATAGGAATTTATTAAGAACTAGCAAGTTTGAATAGGAAAAATGAATAATTTA	5526
QY	1201	GATGATTTAGATGAAGAAATGAATAATCATCAGAAAGATTTATCTGAAGAAAAATATAAA	1260
Db	5527	GATGATTTAGATGAAGAAATGAATAATCATCAGAAAGATTTATCTGAAGAAAAATATAAA	5586
QY	1261	AAAGGAAAGAAATATGAAAAAAACAAAGGATATAATTTTAAACCAATGATPAAAGTTTG	1320
Db	5587	AAAGGAAAGAAATATGAAAAAAACAAAGGATATAATTTTAAACCAATGATPAAAGTTTG	5646
QY	1321	TATGATGACATATTAATAAATAATTAATAAATGATTAAGCAGGTTTAATTAAGAAAAGAAAA	1380
Db	5647	TATGATGACATATTAATAAATAATTAATAAATGATTAAGCAGGTTTAATTAAGAAAAGAAAA	5706
QY	1381	TTCAATAAATCATTTGTTTCATATATTTGACGAGACATGAATAATTTTACAGATCGTGAT	1440
Db	5707	TTCAATAAATCATTTGTTTCATATATTTGACGAGACATGAATAATTTTACAGATCGTGAT	5766
QY	1441	GAGTTATCTGAAGATATACTPAAATATTTTATGAAAACATATAA	1482
Db	5767	GAGTTATCTGAAGATATACTPAAATATTTTATGAAAACATATAA	5808
RESULT 10	AE014834	253001 bp DNA linear INV 07-OCT-2002	
LOCUS	AE014834	Plasmodium falciparum 3D7 chromosome 10 section 6 of 7 of the	
DEFINITION	complete sequence.		
ACCESSION	AE014834	AE014185	
VERSION	AE014834.1	GI:23495173	
KEYWORDS			
SOURCE			
ORGANISM		Plasmodium falciparum 3D7 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
REFERENCE	1 (bases 1 to 253001)		
AUTHORS	Gardner,M.J., Hall,N., Fung,E., White,O., Beriman,M., Hyman,R.W., Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T., James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kye,S., Chan,M.-S., Nene,V., Shalom,S.J., Sub,B., Peterson,J., Anguilo,S., Perlea,M., Allen,J., Selengut,J., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Franholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummins,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.		
TITLE	Genome sequence of the human malaria parasite Plasmodium falciparum		
JOURNAL	Nature 419 (6906), 498-511 (2002)		
PUBMED	12368864		
REFERENCE	2 (bases 1 to 253001)		
AUTHORS	Gardner,M.J.		
TITLE	Direct Submision		
JOURNAL	Submitted (13-SEP-2002) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA		
FEATURES			
source	1..253001 /organism="Plasmodium falciparum 3D7" /mol_type="genomic DNA" /isolate="3D7" /db_xref="taxon:36329" /chromosome="10" complement(79..107) /note="AT rich" /rpt_type=tandem complement(113..136) /rpt_type=tandem /rpt_unit="(a)n"		
repeat_region			
repeat_region			
repeat_region			

[illegible]

```
repeat_region 2808..2966 /rpt_type=tandem
repeat_region 3142..3171 /rpt_type=tandem
repeat_region /rpt_unit="(a)n"
complement(3274)..3347 /note="AT_rich"
repeat_region /rpt_type=tandem
complement(3391)..3437 /rpt_unit="(ta)n"
repeat_region complement(3439)..3503 /rpt_type=tandem
repeat_region /rpt_unit="(a)n"
complement(3500)..3500 /rpt_type=tandem
repeat_region complement(3551)..3683 /rpt_type=tandem
repeat_region /rpt_unit="(ataaa)n"
complement(3710)..3745 /note="AT_rich"
repeat_region /rpt_type=tandem
complement(3763)..3824 /rpt_type=tandem
repeat_region /rpt_unit="(ta)n"
complement(3871)..4025 /note="AT_rich"
repeat_region /rpt_type=tandem
complement(3922)..3987 /rpt_type=tandem
repeat_region /rpt_unit="(ta)n"
complement(4037)..4082 /rpt_type=tandem
repeat_region /rpt_unit="(ataaa)n"
complement(4080)..4082 /rpt_type=tandem
repeat_region complement(4084)..4104 /rpt_type=tandem
repeat_region /rpt_unit="(a)n"
complement(4135)..4181 /rpt_type=tandem
repeat_region /rpt_unit="(ta)n"
complement(4182)..4313 /note="AT_rich"
repeat_region /rpt_type=tandem
complement(4364)..4466 /rpt_type=tandem
repeat_region complement(4457)..4457 /rpt_type=tandem
repeat_region complement(4472)..4472 /rpt_type=tandem
repeat_region /rpt_unit="(ataaa)n"
complement(4512)..4655 /note="AT_rich"
repeat_region /rpt_type=tandem
complement(4670)..4775 /rpt_type=tandem
repeat_region /rpt_unit="(ta)n"
complement(4768)..4768 /rpt_type=tandem
repeat_region complement(4870)..4912 /rpt_type=tandem
repeat_region complement(4870)..4912 /rpt_type=tandem
repeat_region /rpt_unit="(ta)n"
```

```
repeat_region complement(4918)..5000)
/rpt_type=tandem
/rpt_unit="(ataaa)n"
gene
/locus_tag="PF10_0308"
join(<5162..5350,5488..5535,5617..5802,5902..5981,
6099..6162,6254..6349,6463..6540,6727..>8973)
/locus_tag="PF10_0308"
join(5162..5350,5488..5535,5617..5802,5902..5981,
6099..6162,6254..6349,6463..6540,6727..8973)
/locus_tag="PF10_0308"
/codon_start=1
/product="hypothetical protein"

Query Match 94.9%; Score 1406.8; DB 2; Length 253001;
Best Local Similarity 96.8%; Pred. No. 4,2e-151;
Matches 1435; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

1 CAAGAACCAACAAAGCGATCTAGAACAAAGAGACGTCCTAAAGAAAAGTTGCAAGAACAA 60
179846 CAGGCGCACAAACCGATCTAGAACAAAGAGACTAGCTAAAGAAAAGTTACAGGGCCA 179905
61 CAAAGCGATTTAGAACAAAGATAGACTTGCTAAAGAAAAGTTACAAAGACCAAGCGAT 120
179906 CAAAGCGATCTAGAACAAAGAGACTAGCTAAAGAAAAGTTACAGGGCCAAGCGAT 179965
121 TTGAAACAAAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAACCAAGCGATCTAGAACAA 180
179966 TTAGAACAAAGAGACTTGCTAAAGAAAAGTTGCAAGAGCAACAAAGCGATTTAGAACAA 180025
181 GAGAGAGTCGCTAAAGAAAAGTTGCAAGAACAAACCAAGCGATTTAGAACAAAGAGACGT 240
180026 GAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAACCAAGCGATTTAGAACAAAGAGACTA 180085
241 GCTAAAGAAAAGTTGCAAGAACAAACCAAGCGATTTAGAACAAAGATAGACTTGCTAAAGAA 300
180086 GCTAAAGAAAAGTTGCAAGAACAAACCAAGCGATTTAGAACAAAGCACTTGCTAAAGAA 180145
301 AAGTTACAAAGACGCAAGCGATTTAGAACAAAGAGACGTCGCTAAAGAAAAGTTGCCAA 360
180146 AAGTTGCAAGAACCAACCAAGCGATTTAGAACAAAGAGACTAGCTAAAGAAAAGTTACAG 180205
361 GAACCAACAAAGCGATTTAGAACAAAGAGACGTCGCTAAAGAAAAGTTGCAAGAACCAACAA 420
180206 GGGCAGCATAGCGATTTAGAACCAAGCGATTTAGAACAAAGCGTTGCAAGAACCAACAA 180265
421 AGCGATTTAGAACCAAGAGACTTGCTAAAGAAAAGTTGCAAGAACCAACCAAGCGATTTA 480
180266 AGCGATTTAGAACCAAGAGACTTGCTAAAGAAAAGTTGCAAGAACCAACCAAGCGATTTA 180325
481 GAACCAAGAGACGTCGCTAAAGAAAAGTTGCAAGAACCAACCAAGCGATTTAGAACCAAGAG 540
180326 GAACCAAGAGACGTCGCTAAAGAAAAGTTGCAAGAACCAACCAAGCGATTTAGAACCAAGAG 180385
541 AGACGTCGCTAAAGAAAAGTTGCAAGAACCAACCAAGCGATTTAGAACCAAGAGACGTCCT 600
180386 AGACGTCGCTAAAGAAAAGTTGCAAGAACCAACCAAGCGATTTAGAACCAAGAGACGTCCT 180445
601 AAAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACCAAGAGAGCTGATACGAAAAAAA 660
180446 AAAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACCAAGAGAGCTGATACGAAAAAAA 180505
661 AATTTAGAAAAGAAAAGAAACATGAGATATATTTAGCAAGATTTATATGTCGTTTTA 720
180506 AATTTAGAAAAGAAAAGAAACATGAGATATATTTAGCAAGATTTATATGTCGTTTTA 180565
721 GAAATACAGCTATAGAACTTCATCAGAAAAGAAAGCTGATTTATATACCATCA 780
180566 GAAATACAGCTATAGAACTTCATCAGAAAAGAAAGCTGATTTATATACCATCA 180625
781 TCTTCTTTACCTCAGACCAACAGAGGAAATAGTAGATTTCAAGGAAATATCTATATA 840
180626 TCTTCTTTACCTCAGACCAACAGAGGAAATAGTAGATTTCAAGGAAATATCTATATA 180685
```


QY 841 GAAAAACAATAGAGATCTATTACAATAAGTTGAAGAGAGGATATACATATA 900
|
|
|
Db 180866 GAAAAATACAAATAGAGATCTATTACACAAATGTTGAAGAGAGGATATATATA 180745
|
|
|
QY 901 GGACATCTTTGAAGAAAGAGAGATGTTCAATMAAACAGAAACAAAGAGATMAATCT 960
|
|
|
Db 180746 GGACATCTTTGAAGAAAGAGAGATGTTCAATMAAACAGAAACAAAGAGATMAATCT 180805
|
|
|
QY 961 GCTGACATACAAATCATCATTTAGAGACAGTMAATTTTCTGATGTTAATGATTTTCAA 1020
|
|
|
Db 180806 GCTGACATACAAATCATCATTTAGAGACAGTMAATTTTCTGATGTTAATGATTTTCAA 180865
|
|
|
QY 1021 ATTAAGATATAGAGATGAATTAAGTCTGAATTAAGACATTCATTAATAGATGAAGA 1080
|
|
|
Db 180866 ATTAAGATATAGAGATGAATTAAGTCTGAATTAAGACATTCATTAATAGATGAAGA 180925
|
|
|
QY 1081 GAAGATGATGAAGACTTAGACGAATTTAAGCCTATTGTGCAATATGACAAATTCAGAGAT 1140
|
|
|
Db 180926 GAAGATGATGAAGACTTAGACGAATTTAAGCCTATTGTGCAATATGACAAATTCAGAGAT 180985
|
|
|
QY 1141 GAAGAAACATAGAAATTTTATAAGAACTAGAAAGATTTGATAGAGAAAAATGAATTTTA 1200
|
|
|
Db 180986 GAAGAAACATAGAAATTTTATAAGAACTAGAAAGATTTGATAGAGAAAAATGAATTTTA 181045
|
|
|
QY 1201 GATGATTTAGATGAAGAAATAGAAAAATCATCAGAAAGATTTATCTGAAGAAAAATTAATA 1260
|
|
|
Db 181046 GATGATTTAGATGAAGAAATAGAAAAATCATCAGAAAGATTTATCTGAAGAAAAATTAATA 181105
|
|
|
QY 1261 AAAGAAAGAAATATGAAAAAAACAAAGATAATTTTAAACCAATGATMAAGTTTG 1320
|
|
|
Db 181106 AAAGAAAGAAATATGAAAAAAACAAAGATAATTTTAAACCAATGATMAAGTTTG 181165
|
|
|
QY 1321 TATGATGACATATTTAAAAATTTAAAAATGATTAAGCAGTTAATTAAGAAAAAGAAAA 1380
|
|
|
Db 181166 TATGATGACATATTTAAAAATTTAAAAATGATTAAGCAGTTAATTAAGAAAAAGAAAA 181225
|
|
|
QY 1381 TTCATTAATCATTTGTTTATATTTGACGAGACATGAATTTTACAGATCGTGAT 1440
|
|
|
Db 181226 TTCATTAATCATTTGTTTATATTTGACGAGACATGAATTTTACAGATCGTGAT 181285
|
|
|
QY 1441 GAGTATCTGAAGATATACTAAATTTTATGAAACTATA 1482
|
|
|
Db 181286 GAGTATCTGAAGATATACTAAATTTTATGAAACTATA 181327
|
|
|
RESULT 11
A28740 950 bp DNA linear PAT 04-JUN-1995
LOCUS A28740 cDNA for LSA-R-NR protein.
DEFINITION A28740
ACCESSION A28740
VERSION A28740.1 GI:1247512
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 950)
TITLE LIVER-STAGE-SPECIFIC PEPTIDE SEQUENCES OF P. FALCIPARUM BEARING
EPIPTOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES
JOURNAL Patent: WO 9213884-A 29 20-AUG-1992;
FEATURES
Location/Qualifiers
1..950
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 64.1%; Score 950; DB 6; Length 950;
Best Local Similarity 100.0%; Pred. No. 1.8e-98;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AAACGATCTAGAACAAAGAGAGAGCGTGCTAAAGAAAAAGTTGCAAGAACAAAGCGATT 70
|
|
|
Db 1 AAACGATCTAGAACAAAGAGAGAGCGTGCTAAAGAAAAAGTTGCAAGAACAAAGCGATT 60
|
|
|
QY 71 TAGAACAAAGATGACTTGCTTAAAGAAAAGTTTCAAGACAGCAACAGCATTTAGAACAG 130
|
|
|
Db 61 TAGAACAAAGATGACTTGCTTAAAGAAAAGTTTCAAGACAGCAACAGCATTTAGAACAG 120
|
|
|
QY 131 AGAGCTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGAGAGCGTG 190
|
|
|
Db 121 AGAGCTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGAGAGCGTG 180
|
|
|
QY 191 CTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTGCTAAAGAAA 250
|
|
|
Db 181 CTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTGCTAAAGAAA 240
|
|
|
QY 251 AGTTGCAAGAACAAAGCGATTTAGAACAAAGATGACTTGCTTAAAGAAAAGTTTAAAG 310
|
|
|
Db 241 AGTTGCAAGAACAAAGCGATTTAGAACAAAGATGACTTGCTTAAAGAAAAGTTTAAAG 300
|
|
|
QY 311 AGCAGCAAGAGGATTTAGAACAGAGAGAGCGTGCTAAAGAAAAGTTGCAAGAACAA 370
|
|
|
Db 301 AGCAGCAAGAGGATTTAGAACAGAGAGAGCGTGCTAAAGAAAAGTTGCAAGAACAA 360
|
|
|
QY 361 GCGATTTAGAACAGAGAGAGCGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAG 420
|
|
|
Db 431 AACCAAGAGAGCTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 490
|
|
|
QY 421 AACCAAGAGAGCTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 480
|
|
|
Db 491 GACGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGTGCTA 550
|
|
|
QY 481 GACGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGTGCTA 540
|
|
|
Db 551 AAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGTGCTAAAGAAAAGT 610
|
|
|
QY 541 AAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGTGCTAAAGAAAAGT 600
|
|
|
Db 601 TGCAGAGACACAAAGAAATTTAGAACAAAGAGAGCGTGATACGAAAAAAATTTAGAAA 660
|
|
|
QY 611 TGCAGAGACACAAAGAAATTTAGAACAAAGAGAGCGTGATACGAAAAAAATTTAGAAA 670
|
|
|
Db 661 GAAAAAAGGAAACATGAGATATATTAGAGAGAGATTTATATGTCGTTTGAATATACAG 720
|
|
|
QY 721 CTATAGAACTTCCATCAGAAAAATGAAACGTGATATTATATACCATCATCTTCTTAC 790
|
|
|
Db 791 CTATAGAACTTCCATCAGAAAAATGAAACGTGATATTATATACCATCATCTTCTTAC 780
|
|
|
QY 791 CTCAGACACACAGAGGAATAGTAGAGATTCACAGAAAAATCTATTAATGAAAAAACAA 850
|
|
|
Db 841 CTCAGACACACAGAGGAATAGTAGAGATTCACAGAAAAATCTATTAATGAAAAAACAA 840
|
|
|
QY 851 ATAGAGATCTATTACAAACAATGTTGAAGGCGAAGGAGATATACATAAGAGACATCTTG 910
|
|
|
Db 901 ATAGAGATCTATTACAAACAATGTTGAAGGCGAAGGAGATATACATAAGAGACATCTTG 900
|
|
|
RESULT 12
AR162961 950 bp DNA linear PAT 17-OCT-2001
LOCUS AR162961
DEFINITION Sequence 32 from patent US 6270771.
ACCESSION AR162961
VERSION AR162961.1 GI:16233425
KEYWORDS

SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 950)
AUTHORS Guerin-Marchand, C. and Druilhe, P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 620771-A 32 07-AUG-2001;
FEATURES
 source
 1..950
 /organism="unknown"
 /mol_type="unassigned DNA"
ORIGIN
Query Match 64.1%; Score 950; DB 6; Length 950;
Best Local Similarity 100.0%; Pred. No. 1.8e-98;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 AAAGCGATCTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 70
DB 1 AAAGCGATCTAGAACAAAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 60
QY 71 TAGAACAAAGATAGACTGTCTAAAGAAAAGTTACAAAGAGAGCAAGCGATTAGAACAG 130
DB 61 TAGAACAAAGATAGACTGTCTAAAGAAAAGTTACAAAGAGAGCAAGCGATTAGAACAG 120
QY 131 AGAGACTTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAAAGAGAGCGTG 190
DB 121 AGAGACTTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAAAGAGAGCGTG 180
QY 191 CTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAA 250
DB 181 CTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAA 240
QY 251 AGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTAGAAC 310
DB 241 AGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTAGAAC 300
QY 311 AGCAGCAAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 370
DB 301 AGCAGCAAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 360
QY 371 GCCATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAG 430
DB 361 GCCATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAG 420
QY 431 AACAAAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAG 490
DB 421 AACAAAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAG 480
QY 491 GACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAGAGCGTGCTA 550
DB 481 GACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAGAGCGTGCTA 540
QY 551 AAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAAGT 610
DB 541 AAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAAGT 600
QY 611 TGCAAGAGCAGCAAGAGATTAGAACAAAGAGAGCGATAGCAAGAAAAGTTAGAAA 670
DB 601 TGCAAGAGCAGCAAGAGATTAGAACAAAGAGAGCGATAGCAAGAAAAGTTAGAAA 660
QY 671 GAAAAAAGCAATGAGAGATATTAGCAGAGATTATATGTCGTTTAGAAAATACAG 730
DB 661 GAAAAAAGCAATGAGAGATATTAGCAGAGATTATATGTCGTTTAGAAAATACAG 720
QY 731 CTATAGAACTTCATCAGAAAATGAAAGTGATATTATACCATCAATCTCTTTAC 790
DB 721 CTATAGAACTTCATCAGAAAATGAAAGTGATATTATATACCATCAATCTCTTTAC 780
QY 791 CTCAGAGACAGAGAGGAAATAGTAGAGATTCAGAGAAATATCTATAATAGAAAACAA 850
DB 781 CTCAGAGACAGAGAGGAAATAGTAGAGATTCAGAGAAATATCTATAATAGAAAACAA 840

QY 851 ATAGAGAAATCTATTACAACAATGTTGAAGAGAGGATATATATAAGACATCTTG 910
DB 841 ATAGAGAAATCTATTACAACAATGTTGAAGAGAGGATATATATAAGACATCTTG 900
QY 911 AAGAAAAGAAAGATGGTTCAATTAACAGAACAAAAGAGATTAATCT 960
DB 901 AAGAAAAGAAAGATGGTTCAATTAACAGAACAAAAGAGATTAATCT 950
RESULT 13
AR178250 950 bp DNA linear PAT 20-APR-2002
LOCUS Sequence 32 from patent US 6319502.
DEFINITION AR178250
ACCESSION AR178250
VERSION AR178250.1 GI:20219388
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 950)
AUTHORS Guerin-Marchand, C. and Druilhe, P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6319502-A 32 20-NOV-2001;
FEATURES
 source
 1..950
 /organism="unknown"
 /mol_type="unassigned DNA"
ORIGIN
Query Match 64.1%; Score 950; DB 6; Length 950;
Best Local Similarity 100.0%; Pred. No. 1.8e-98;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 AAAGCGATCTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 70
DB 1 AAAGCGATCTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 60
QY 71 TAGAACAAAGATAGACTGTCTAAAGAAAAGTTACAAAGAGAGCAAGCGATTAGAACAG 130
DB 61 TAGAACAAAGATAGACTGTCTAAAGAAAAGTTACAAAGAGAGCAAGCGATTAGAACAG 120
QY 131 AGAGACTTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAAAGAGAGCGTG 190
DB 121 AGAGACTTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAAAGAGAGCGTG 180
QY 191 CTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAA 250
DB 181 CTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAA 240
QY 251 AGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTAGAAC 310
DB 241 AGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTAGAAC 300
QY 311 AGCAGCAAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 370
DB 301 AGCAGCAAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 360
QY 371 GCCATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAG 430
DB 361 GCCATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAG 420
QY 431 AACAAAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAG 490
DB 421 AACAAAGAGAGCTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAG 480
QY 491 GACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTGCTA 550
DB 481 GACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTGCTA 540
QY 551 AAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAGAGCGTCTAAAGAAAAGT 610

Db	541	AAGAAAAGTTGCACGAAACAACAAGGCCATTTAGAACCAAGAGACGGTGCTAAGAAAAAGT	600
Oy	611	TGCAGAAGCCACCAAGAGATTGTAGACCAAGGAAGCGTATACGAAAAAAAAATTAGAAA	670
Db	601	TGCAAGAGCACCAAAGATTTAGAACCAAGGAAGCGCTGATACGAAAAAAATTAGAAA	660
Oy	671	GAAAAAAGCAACATCGAGATATATATTAGACAGAGATTTATATAGTCGTTTGAATTACCAG	730
Db	661	GAAAAAAGCAACATCGAGATATATATTAGACAGAGATTTATATAGTCGTTTGAATTACCAG	720
Oy	721	CTATGAACTTCATCAGAAAAATGCAACGTGATATTATATCCACATCAATCTCTTTAC	790
Db	721	CTATGAACTTCATCAGAAAAATGCAACGTGATATTATATCCACATCAATCTCTTTAC	780
Oy	791	CTCAGACCAACAGAGGAGATAGTAGATTCGAAAGAAATCTATATATAGAAAAACAA	850
Db	781	CTCAGACCAACAGAGGAGATAGTAGATTCGAAAGAAATCTATATATAGAAAAACAA	840
Oy	851	ATAGAGATCTATTTCACCAAAATGTTGAAGACCAAGGATATATCATTAAGACATCTTG	910
Db	841	ATAGAGATCTATTTCACCAAAATGTTGAAGACCAAGGATATATCATTAAGACATCTTG	900
Oy	911	AAGAAAAAGATGCTTCATTAATAACAGAACCAAAAAGATTAATCTT	960
Db	901	AAGAAAAAGATGCTTCATTAATAACAGAACCAAAAAGATTAATCTT	950
RESULT 14			
PfPRLSAC		843 bp DNA linear INV 19-AUG-1995	
LOCUS		Plasmodium falciparum (clones KEN1-21b, KEN1-40h, KEN1-56b,	
DEFINITION		PNM1-28a, PNM1-28c, BRA1-19e) liver stage-specific antigen 1	
ACCESSION		(USA-1) gene, 3' end of cdf.	
VERSION		L40885	
KEYWORDS		L40885.1 GI:725278	
REFERENCE		exoerythrocytic; liver stage-specific antigen; liver stage-specific	
SOURCE		antigen 1; malaria.	
ORGANISM		Plasmodium falciparum (malaria parasite P. falciparum)	
ISOLATE		Plasmodium falciparum	
REFERENCE		Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
AUTHORS		1 (bases 1 to 843)	
TITLE		Yang,C., Shi,Y.P., Udhayakumar,V., Alpers,M.P., Povoa,M.M.,	
JOURNAL		Hawley,W.A., Collins,W.E. and Lal,A.A.	
PUBMED		Sequence variations in the non-repetitive regions of the liver	
COMMENT		stage-specific antigen-1 (USA-1) of Plasmodium falciparum from	
		field isolates	
		Mol. Biochem. Parasitol. 71 (2), 291-294 (1995)	
		7477115	
		Original source text: Plasmodium falciparum (individual isolate	
		Kenyan) (clone: KEN1-56b) blood stage parasite DNA; Plasmodium	
		falciparum (individual isolate Kenyan) (clone: PNM1-28a) blood	
		stage parasite DNA; Plasmodium falciparum (individual isolate	
		Kenyan) (clone: PNM1-28c) blood stage parasite DNA; Plasmodium	
		falciparum (individual isolate Kenyan) (clone: KEN1-21b) blood	
		stage parasite DNA; Plasmodium falciparum (individual isolate	
		Kenyan) (clone: KEN1-40h) blood stage parasite DNA; Plasmodium	
		falciparum (individual isolate Kenyan) (clone: BRA1-19e) blood	
		stage parasite DNA.	
		Clone name abbreviations: KEN for Kenya, PNG for Papua New Guinea	
		and BRA for Brazil. For each clone, the number following the KEN,	
		PNG or BRA represents group number, the number following the '-'	
		denotes the patient's blood sample number and the lower case letter	
		represents recombinant clones from the corresponding isolates.	
FEATURES			
source		Location/Qualifiers	
		1..843	
		/organism="Plasmodium falciparum"	
		/mol_type="genomic DNA"	
		/isolate="Kenyan"	
		/db_xref="taxon:5833"	
		/clone="KEN1-56b"	
		/clone="PNM1-28a"	
		/clone="PNM1-28c"	
		/clone="KEN1-21b"	

	gene	/clone="KEN1-40h"
	CDS	/dev_stage="blood stage parasite"
		1..843
		/gene="LSA-1"
		<1..843
		/gene="LSA-1"
		/codon_start=1
		/product="liver stage-specific antigen 1"
		/protein_id="AA59231.1"
		/db_xref="GI:950664"
		/translation="RKADTKKNLERKKEHGDILAEIDYGRLEIPALIELPSENERGYV PHOSLLODRNGNRSDSKSEISIEKTRESITTVTEGSRDIDGHLEEKQSGIKPEP KEDSAIIOHHLETJVNISPVNDROI SKYDESI SAEXDDSI IDDEEDDEDIDEPFPI QYDFDQBEENIGIYLEDLERKENIDDDDEGISSEBELSERIKKGKYERTKD NFKVDNSLYDEHIKKYKNDQYNKEKEFIKS LFI FIDGNELIQIVDELSEDITTK FMKL"
	variation	228
		/gene="LSA-1"
		/note="Clone PNGI-28a"
		/replace="T"
ORIGIN		
Query Match	56.9%; Score 843; DB 2; Length 843;	
Best Local Similarity	100.0%; Pred. No. 2,3e-86;	
Matches 843; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Oy	640 AGAAGCGCTGATACGAAAAAAAATTGGAAAGAAAAAGAACAATGAGATATATTACA	699
Dd	1 AGGAAGCTGATACGAAAAAAAATTGGAAAGAAAAAGAACAATGAGATATATTACA	60
Oy	700 GAGATTTATATGTGCTGTTTGAAATACACACTATAGAACCTTCATCGAAAAATGAACT	759
Dd	61 GAGATTTATATGTGCTGTTTGAAATACCACTATAGAACTTCATCGAAAAATGAACT	120
Oy	760 GGATTTATATACACATCATCTCTCTTACTCAGACAACAGAGGATATGAGAT	819
Dd	121 GGATTTATATACACATCATCTCTTACTCAGACAACAGAGGATATGAGAT	180
Oy	820 TCCAAGAAATATCTATATAGAAAAAACAATAGAGATCTATTACAACAATGTTGA	879
Dd	181 TCCAAGAAATATCTATATAGAAAAAACAATAGAGATCTATTACAACAATGTTGA	240
Oy	880 GCACGAGGATATACATAAAGACATCTTGAGAAAAAGAAAGATGTTCAATAAACCA	939
Dd	241 GCACGAGGATATACATAAAGACATCTTGAGAAAAAGAAAGATGTTCAATAAACCA	300
Oy	940 GAACAAAAAGAGTAATCTGCCTGACATCAAAATCATCAATTAGACAGTAATATT	999
Dd	301 GAACAAAAAGAGTAATCTGCCTGACATCAAAATCATCAATTAGACAGTAATATT	360
Oy	1000 TCTGATGTTAATGATTTTCAAATAGTAGATGAGATGAAATTAAGCTGATATGAC	1053
Dd	361 TCTGATGTTAATGATTTTCAAATAGTAGATGAGATGAAATTAAGCTGATATGAC	420
Oy	1060 GATTCATTAATAGATGAAGAAAGAGATGAGAAGCTTAGAGCAATTTAAGCCATTGTG	1113
Dd	421 GATTCATTAATAGATGAAGAAAGAGATGAGAAGCTTAGAGCAATTTAAGCCATTGTG	480
Oy	1120 CAATRTGCAATTTCCAAAGATGAAGAAAACTAGAAATTTATTAAGAACTAGAAATTTG	1177
Dd	481 CAATRTGCAATTTCCAAAGATGAAGAAAACTAGAAATTTATTAAGAACTAGAAATTTG	540
Oy	1180 ATAGAGAAAAATGAAAAATTTAGATGATTAGATGAAGAAATGAAAAATCATCAGAGAA	1233
Dd	541 ATAGAGAAAAATGAAAAATTTAGATGATTAGATGAAGAAATGAAAAATCATCAGAGAA	600
Oy	1240 TTATCTGAAGAAAAATAAAAAAGGAAGAAATATGAAAAAACAAGATATATATTT	1293
Dd	601 TTATCTGAAGAAAAATAAAAAAGGAAGAAATATGAAAAAACAAGATATATATTT	660
Oy	1300 AAACCAATGATAAAAAGTTGTATGATGAGCATATTAATAATTAATAATGATTAACAG	1355

Db 661 AAACCAATGATTAAGTTTGTATGATGACATATTAATAAATATTAATAATGATAACAG 720

Qy 1360 GTTAATAAGAAAAAGAAAAATTCATAAAATCATTTGTTTCATATATTTGACGAGACAT 1419

Db 721 GTTAATAAGAAAAAGAAAAATTCATAAAATCATTTGTTTCATATATTTGACGAGACAT 780

Qy 1420 GAAATTTTACGATCGTGAGTATCTGAAATATTAATAATTTTATGAACTA 1479

Db 781 GAAATTTTACGATCGTGAGTATCTGAAATATTAATAATTTTATGAACTA 840

Qy 1480 TAA 1482

Db 841 TAA 843

RESULT 15

PFALSB

LOCUS

DEFINITION 843 bp DNA linear INV 19-AUG-1995

Plasmodium falciparum (clones KEN0-11a, -11b, -37d, -39a, -44e, -51a, -53c, -53d, -56a, PNG0-5h, -7a, -8f, -19a, -26b, -86h, -86i, BRA0-5a) liver stage-specific antigen 1 (LSA-1) gene, 3' end of cds.

ACCSSION

VERSION L40884.1 GI:725277

KEYWORDS exoerythrocytic; liver stage-specific antigen; liver stage-specific antigen 1; malaria.

SOURCE Plasmodium falciparum (malaria parasite P. falciparum)

ORGANISM Plasmodium falciparum Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 843)

AUTHORS Yang, C., Shi, Y. P., Udhayakumar, V., Alpers, M. P., Povos, M. M., Hawley, W. A., Collins, W. E. and Lal, A. A.

TITLE Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates

JOURNAL Mol. Biochem. Parasitol. 71 (2), 291-294 (1995)

PIRMBED 747715

COMMENT Original source text: Plasmodium falciparum (individual isolate Kenyan) (clone: KEN0-44e) DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN0-51a) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-7a) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-5h) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN0-11a) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN0-11b) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN0-37d) DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN0-39a) DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-8f) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-19a) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-26b) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-86h) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-86i) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN0-53c) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN0-53d) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN0-56a) blood stage parasite DNA; and Plasmodium falciparum (individual isolate Kenyan) (clone: BRA0-5a) blood stage parasite DNA.

FEATURES

Source

1..843

Location/Qualifiers

1..843

organism="Plasmodium falciparum"

/mol_type="genomic DNA"

/isolate="Kenyan"

/db_xref="taxon:5833"

/clone="KEN0-51a"

/clone="KEN0-44e"

/clone="KEN0-7a"

/clone="PNG0-5h"

/clone="KEN0-11a"

/clone="KEN0-11b"

/clone="KEN0-37d"

/clone="KEN0-39a"

/clone="PNG0-8f"

/clone="PNG0-19a"

/clone="PNG0-26b"

/clone="PNG0-86h"

/clone="PNG0-86i"

/clone="KEN0-53c"

/clone="KEN0-53d"

/clone="KEN0-56a"

/clone="BRA0-5a"

/dev_stage="blood stage parasite"

1..843

/gene="LSA-1"

<1..843

/gene="LSA-1"

/codon_start=1

/product="liver stage-specific antigen 1"

/protein_id="AAB59230.1"

/db_xref="GI:950863"

/translation="RKADYTKNLKERKKGVDVLAIEDYGRLEIPALEPSENERGYI PHOSLIPQDNRGNSRDSKEISIEKTNRESITTNVGRDRIHGHLEEKDGSIKPEQ KEDKSADIQHTLETYNI SPVNDFOISKYEISAEYDSDLIBEDBEDDUEKPYV QYNPDFENIGIKYKELEDLIEKNENDLDDEGIEKSESEELSEEKIKKGRYEKTKN NRPNDNLSYDEHIKKYKNDKQVKEKEKIKSLFPHFDGDNELIQVLSLSDITKY FMKL"

60

/gene="LSA-1"

/note="clone BRA0-5a"

/replace="g"

variation

ORIGIN

Query Match 56.8%; Score 841.4; DB 2; Length 843;

Best Local Similarity 99.9%; Pred. No. 3.5e-86;

Matches 842; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 640 AGCAAGCTGATACGAAAAAATTTAGAAAAAGAAAAAGAACTGAGATATTTTGA 699

Db 1 AGCAAGCTGATACGAAAAAATTTAGAAAAAGAAAAAGAACTGAGATATTTGCA 60

Qy 700 GAGGATTTATGATGCTGTTAGAAATACACGCTATAGAACTTCATCAGAAAAATGAACT 759

Db 61 GAGGATTTATGATGCTGTTAGAAATACACGCTATAGAACTTCATCAGAAAAATGAACT 120

Qy 760 GGATATTATATACACATCAATCTCTTACCTCAGACAACAGAGGAAATAGTAGAT 819

Db 121 GGATATTATATACACATCAATCTCTTACCTCAGACAACAGAGGAAATAGTAGAT 180

Qy 820 TCCAAGAAATATTTATTAATGAAAAAACAATGAGAACTTATTAACAACAAGTTGAA 879

Db 181 TCCAAGAAATATTTATTAATGAAAAAACAATGAGAACTTATTAACAACAAGTTGAA 240

Qy 880 GGAGGAAGGATATACATTAAGAGCATCTTGAAGAAAGAAAGATGGTTCAATTAACA 939

Db 241 GGAGGAAGGATATACATTAAGAGCATCTTGAAGAAAGAAAGATGGTTCAATTAACA 300

Qy 940 GAACAAAAAGAGTAATATGCTGACATACAAATCATACATTAGAGACAGTAATATT 999

Db 301 GAACAAAAAGAGTAATATGCTGACATACAAATCATACATTAGAGACAGTAATATT 360

Qy 1000 TCTGATGTTAATGATTTTCAAAATAGTAAGTATGAGATGAATTAATGCTGGAATAGAC 1059

Db 361 TCTGATGTTAATGATTTTCAAAATAGTAAGTATGAGATGAATTAATGCTGGAATAGAC 420

Qy 1060 GATTCATTATATAGTGAAGAAAGATGATGAACCTTAGACGATTTAAGCCTATTGTG 1119

```

Db      421 GATTCATTAAATGATGAGAGAGATGATGAAGACTTAGACGAATTTAAGCCATATGTG 480
QY      1120 CAATGTGACAAATTTCCAGATGAGAAAAACATAGCAATTTATAAAGAACTAGAAATTTG 1179
Db      481 CAATGTGACAAATTTCCAGATGAGAAAAACATAGCAATTTATAAAGAACTAGAAATTTG 540
QY      1180 ATAGAGAAAAATGAAAAATTTAGATGATTTAGATGAAGGAAATGAAAAATCATCAGAGAA 1239
Db      541 ATAGAGAAAAATGAAAAATTTAGATGATTTAGATGAAGGAAATGAAAAATCATCAGAGAA 600
QY      1240 TTATCTGAAGAAAAATATAAAAAAGGAAAGAAATATGAAAAAACAAGGATATATATTT 1299
Db      601 TTATCTGAAGAAAAATATAAAAAAGGAAAGAAATATGAAAAAACAAGGATATATATTT 660
QY      1300 AAACCAATGATAAAGTTGTATGATGACATATTAATAAATATTAATAAATGATAAGCAG 1359
Db      661 AAACCAATGATAAAGTTGTATGATGACATATTAATAAATATTAATAAATGATAAGCAG 720
QY      1360 GTTATTAAGGAAAGAAAAATTCATTAATCATTTGTTTCATATATTTGACGAGACAAT 1419
Db      721 GTTATTAAGGAAAGAAAAATTCATTAATCATTTGTTTCATATATTTGACGAGACAAT 780
QY      1420 GAAATTTTACAGATCGTGATGAGTTATCTGAAGATATTAATAAATTTTATGAAACTA 1479
Db      781 GAAATTTTACAGATCGTGATGAGTTATCTGAAGATATTAATAAATTTTATGAAACTA 840
QY      1480 TAA 1482
Db      841 TAA 843

```

Search completed: December 7, 2005, 01:48:02
 Job time : 7831 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: December 7, 2005, 06:11:21 ; Search time 33 Seconds

(Without alignments)
8642.015 Million cell updates/sec

Title: US-09-837-344-41
Perfect score: 2486
Sequence: 1 CAAGACACACAAAGCGATCT.....AATTTTATGAAACTATTA 1482

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 28346 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+nzp.model -DEV=xjh
-Q/cgn2.1/USPTO.spool/h/US09837344/runac.05122005.133407.29473/app_query.fasta.1.1671
-DB=PIR -OPMT=fastan -SUFFIX=rrp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -STRAT=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-DOCLIN=200 -THR_SCORE=spc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09837344 @CGN 1 1 40 @runac.05122005.133407.29473 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESUBSTRY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARR_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: pir80.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2406.5	96.8	1909	2	A45592 liver stage antigen
2	380.5	15.3	1407	1	A28589 trichohyalin - rab
3	380.5	15.3	1549	1	A40691 trichohyalin - she
4	376	15.1	1898	1	A45973 trichohyalin - hum
5	356	14.3	1017	2	T15598 hypothetical prote
6	353.5	14.2	737	2	T15597 hypothetical prote
7	333.5	13.4	1192	2	A71623 probable secreted
8	326.5	13.1	1979	1	S03166 myosin heavy chain
9	323.5	13.0	1039	2	S18199 myosin heavy chain
10	321	12.9	771	1	A33430 h-caldesmon - chic
11	319	12.8	1938	2	JCS421 smooch muscle myos
12	319	12.8	1972	2	JCS420 smooch muscle myos
13	318.5	12.8	1961	1	A61231 myosin heavy chain
14	317.5	12.8	1166	2	H86341 hypothetical prote

15	317.5	12.8	1937	2	I38055 myosin heavy chain
16	317.5	12.8	1992	2	A47297 myosin heavy chain
17	317	12.8	1432	2	B85431 trichohyalin like
18	316	12.7	1940	2	A29320 myosin heavy chain
19	314	12.6	1938	2	A59293 skeletal myosin he
20	314	12.6	1976	2	A59252 myosin heavy chain
21	313.5	12.6	1738	2	T14867 interaptin - slime
22	312.5	12.6	1959	1	A33977 myosin heavy chain
23	312	12.6	2007	1	B43402 myosin heavy chain
24	311.5	12.5	2139	2	T18296 myosin heavy chain
25	310.5	12.5	665	2	B71609 hypothetical prote
26	310.5	12.5	678	2	A54514 glutamic acid-rich
27	310	12.5	1963	1	MMKW myosin heavy chain
28	310	12.5	2116	2	A26655 myosin heavy chain
29	309.5	12.4	1956	2	T16416 hypothetical prote
30	309	12.4	1931	2	A59234 slow myosin heavy
31	309	12.4	1972	1	A41604 myosin heavy chain
32	307.5	12.4	1701	2	T09127 probable erythrocy
33	307	12.3	746	2	T47237 myosin II heavy ch
34	306.5	12.3	522	2	C96608 hypothetical prote
35	306.5	12.3	1780	2	S67593 transport protein
36	306	12.3	1999	1	S21801 myosin heavy chain
37	305.5	12.3	2017	1	A36014 myosin heavy chain
38	305.5	12.3	2057	2	S61477 myosin II heavy ch
39	305	12.3	1964	2	A59282 nonmuscle myosin I
40	304.5	12.2	1939	1	A46762 myosin alpha heavy
41	304	12.2	451	2	G70241 hypothetical prote
42	303.5	12.2	1938	2	T49464 alpha cardiac myos
43	303.5	12.2	1939	2	T48175 myosin heavy chain
44	303	12.2	1526	2	A45605 mature-parasite-in
45	302.5	12.2	1130	2	T34081 hypothetical prote

ALIGNMENTS

RESULT 1
A45592
liver stage antigen LSA-1 - malaria parasite (Plasmodium falciparum)
C/Dates: Plasmodium falciparum
C/Date: 22-Nov-1993 #sequence revision 02-Dec-1994 #text change 09-Jul-2004
C/Accession: S24597; A45592; S29393; S34842; B45592; D45592
R/Zhu, J.; Hollingdale, M.
Submitted to the EMBL Data Library, November 1990
A/Reference number: S24597
A/Accession: S24597
A/Molecule type: DNA
A/Residues: 1-1909 <ZHU>
A/Cross-references: UNIPROT:Q25893; UNIPARC:UPI000081753; EMBL:X56203; NID:g9915; PID:
R/Zhu, J.; Hollingdale, M.R.
Mol. Biochem. Parasitol. 48, 223-226, 1991
A/Title: Structure of Plasmodium falciparum liver stage antigen-1.
A/Reference number: A45592; MUID:92107224; PMID:1840628
A/Accession: A45592
A/Molecule type: DNA
A/Residues: 1-195,638-668;1165-1215;1590-1909 <ZH2>
A/Cross-references: UNIPARC:UPI0000177378; UNIPARC:UPI0000177379; UNIPARC:UPI000017737A
A/Note: Sequence extracted from NCBI backbone (NCBIN:75010, NCBIN:75012, NCBIN:75014, N
R/Guerin-Marchand, C.; Druilhe, P.; Galey, B.; Londono, A.; Pataraopitkul, J.; Beaudoin
Nature 329, 164-167, 1987
A/Title: A liver-stage-specific antigen of Plasmodium falciparum characterized by gene
A/Reference number: S29393; MUID:87315391; PMID:3306406
A/Accession: S29393
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 323-387 <GUB1>
A/Cross-references: UNIPARC:UPI000017737C; EMBL:M28266
R/Guerin-Marchand, C.; Druilhe, P.; Galey, B.; Londono, A.; Pataraopitkul, J.; Beaudoin
Submitted to the EMBL Data Library, April 1992
A/Description: a liver-stage-specific antigen of Plasmodium falciparum characterized by
A/Reference number: S34842
A/Accession: S34842
A/Status: preliminary
A/Molecule type: DNA

A:Residues: 323-381, 'HKAI', <GUE2>
A:Cross-references: UNIPARC:UPI000017737D; EMBL:M28266
A:Note: difference at carboxyl end due to frameshift error
C:Comment: This protein is found as flocculent material in the parasitophorous vacuole.
C:Superfamily: trichohyalin; calmodulin repeat homology
C:Keywords: EF hand
F:154-1629/Region: 17-residue repeats (A-K-E-K-L-Q-E-Q-S-D-L-E-Q-E-R-R)

Alignment Scores:

Align. No.:	2, 57e-119	Length:	1909
Score:	2406.50	Matches:	487
Percent Similarity:	89.89%	Conservative:	2
Best Local Similarity:	89.52%	Mismatches:	4
Query Match:	96.80%	Indels:	51
DB:	2	Gaps:	2

US-09-837-344-41 (1-1482) x A45592 (1-1909)

```
QY      1 CAAGAACAAACGATCTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCAGAACAA 60
      |||
Db     1366 GInGInGInGInSerAspLeuGInGInGInGInGInGInGInGInGInGInGIn 1385
      |||
QY      61 CAAAGCGATTTGACACAGATGACTT-----
      |||
Db     1386 GInSerAspLeuGInGInAspArgLeuAlaLysGInLysLeuGInGInGInGInGIn 1405
      |||
QY      88 -----GCTAAAGAAAAGTTACAAAGACGACAAAGCGATTAGAACAA 129
      |||
Db     1406 GInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1425
      |||
QY     130 GAGAGACTTGCTAAAGAAAAGTTGCAGAACAAACGAGCTCTAGAACAGAGACGT 189
      |||
Db     1426 GInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1445
      |||
QY     130 GCTAAAGAAAAGTTGCAGAACAAACGAGCTCTAGAACAGAGACGCTCTAAAGAA 249
      |||
Db     1446 AlaLysGInLysLeuGInGInGInGInGInGInGInGInGInGInGInGInGIn 1465
      |||
QY     250 AAGTTGCAGAACAAACGATTTAGAACAAAGATGAGCTCTCTAAAGAAAAGTTACAA 309
      |||
Db     1466 LysLeuGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1485
      |||
QY     310 GAGCAGCAAAAGCGATTTAGAACAGAGACGCTCTAAAGAAAAGTTGCAGAACAA 369
      |||
Db     1486 GInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1505
      |||
QY     370 AGCGATTTGAAACAAAGAGACGCTCTAAAGAAAAGTTGCAGAACAAACGAGCTTTA 429
      |||
Db     1506 SerAspLeuGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1525
      |||
QY     430 GAAACAAGAGAGACTTGCTAAAGAAAAGTTGCAGAACAAACGAGATTTAGAACAGAG 489
      |||
Db     1526 GInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1545
      |||
QY     490 AGACGTGCTAAAGAAAAGTTGCAGAACAAACGAGATTTAGAACAGAGACGCTGCT 549
      |||
Db     1546 ArgArgAlaLysGInLysLeuGInGInGInGInGInGInGInGInGInGInGInGIn 1565
      |||
QY     550 AAAAGAAAAGTTGCAGAACAAACGAGATTTAGAACAGAGACGCT----- 597
      |||
Db     1566 LysGInLysLeuGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1585
      |||
QY     597 -----
      |||
Db     1586 GInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1605
      |||
QY     598 -----GCTAAAGAAAAGTTGCAGAACAGACGACGACAAAGA 627
      |||
Db     1606 GInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1625
      |||
QY     628 GATTTCGAAACAAAGAGCTGATTCGAAAGAAAAGTTTCGAAAGAAAAGAGACATGGA 687
      |||
Db     1626 AspLeuGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1645
```

```
QY     688 GATATATTAGAGAGATTTATATGTCGTTAGAAAATACAGCTATAGAACTTTCATCA 747
      |||
Db     1646 AspValLeuAlaGInAspLeuTyGInArgLeuGInLeuProAlaLeuLeuProSer 1665
      |||
QY     748 GAAATGACGCTGATATTATATACACATCAATCTTTTACCTCAGACACAGAGGG 807
      |||
Db     1666 GInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1685
      |||
QY     808 AATGTTAGAGTTCCAGAGAAATATCTTAATAGAAAACAAATAGAGAACTCTTTCA 867
      |||
Db     1686 AsnSerArgAspSerLysGInLeuSerLeuLeuGInLysTrpAsnArgGInSerLeuThr 1705
      |||
QY     868 ACAATGTTGAGAGCAGAGGAGATATACATAAGGACATCTTGAGAAAAGAGAGATGCT 927
      |||
Db     1706 ThrAsnValGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1725
      |||
QY     928 TCAATTAACCAAGAACAAAGAAATATCTGCTGACATACAAATCATTCATTAGAG 987
      |||
Db     1726 SerLeuLysProGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1745
      |||
QY     988 ACAGTAAATATTTTCGATGTTAATGATTTTCAAAATAGTAAGGATGAAATAGT 1047
      |||
Db     1746 ThrValAsnLeuSerAspValAsnAspGlnLeuSerLysTyGInAspGInLeuSer 1765
      |||
QY     1048 GCTGATATGACGATTCATTAATAGATGAGAACAGAGATGATGAGACTTACGCAATTT 1107
      |||
Db     1766 AlaGInTyArgAspSerLeuLeuAspGInGInGInGInGInGInGInGInGInGIn 1785
      |||
QY     1108 AAGCTTATTTGCAATATGACAAATTTCCAGATGAGAAAACATRGAAATTTATTAAGAA 1167
      |||
Db     1786 LysProLeuValGInTyArgAspAsnGlnGInGInGInGInGInGInGInGInGIn 1805
      |||
QY     1168 CTAGAGATTTGATGAGAAAATGAAAATTTAGTGATTTAGAGGAATGAAATAA 1227
      |||
Db     1806 LeuGInAspLeuLeuGInLysAsnGInLysAsnLeuAspLeuAspGInGInGInGIn 1825
      |||
QY     1228 TCATCAGAGAAATTTATCTGAGAGAAAATTAATAAGAAAAGAAATATGAAAACAAAG 1287
      |||
Db     1826 SerSerGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1845
      |||
QY     1288 GATATATATTTTAAACCAATGATTAAGATTTGATGAGCATTTATTAATAATATAA 1347
      |||
Db     1846 AspAsnAsnPheLysProAsnAspLysSerLeuTyArgAspGInGInGInGInGIn 1865
      |||
QY     1348 AATGATTAAGCAGTTAAATAGAAAAGAAAATTCATTAATCATTTTCATATATT 1407
      |||
Db     1866 AsnAspLysGInValAsnLysGInLysPheLeuLysSerLeuPheLeuLeuPhe 1885
      |||
QY     1408 GACGAGACAATGAAATTTTACAGATCGTGATGAGTTATCTGAAGATATACTAAATAT 1467
      |||
Db     1886 AspGInAspAsnGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1905
      |||
QY     1468 TTTATGAAACTA 1479
      |||
Db     1906 PheMetLysLeu 1909
      |||
```

RESULT 2

S28583
trichohyalin - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 12-Mar-1993 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C:Accession: S28589
R:Fieler, M.J.; Rogers, G.B.
submitted to the EMBL Data Library, December 1992
A:Description: Examination of the gene encoding rabbit trichohyalin.
A:Reference number: S28589
A:Accession: S28589
A:Molecule type: DNA
A:Residues: 1-1407 <FIE>
A:Cross-references: UNIPROT:P37709; UNIPARC:UPI000013738C; EMBL:Z19092; NID:G1746; PIN:
C:Comment: Trichohyalin is a protein of the medulla of the hair and of the inner root st
Covalent modifications to this protein include conversion of arginine to citrulline and

A>Note: sequence extracted from NCBI backbone (NCBIP:132511)
R.Fietz, M.J.: Preeland, R.B.: Rogers, G.E.
J.Cell Biol. 110, 427-436, 1990
A>Title: The cDNA-deduced amino acid sequence for trichohyalin, a differentiation marker
A:Reference number: A34209; M0ID:90130632; PMID:2298812
A:Accession: A34209
A:Molecule type: mRNA
A:Residues: 1016-1151, 1205-1257, 1281-1398, 'G', 1400-1549 <F12>
A:Cross-references: UNIPARC:UPI000016CSF4; GB:X51695; NID:CA35992.1; PID:gl
C:Comment: Trichohyalin is a protein of the medulla of the hair and of the inner root sheath
Covalent modifications to this protein include conversion of arginine to citrulline and
C:Geneid:68
A:introns: 46/3
A>Note: single copy gene
C:Superfamily: trichohyalin; calmodulin repeat homology
C:Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat
F/49-81/Domain: calmodulin repeat homology <EF2>
F/387-851/Region: 28-residue repeats
F/886-1519/Region: 23-residue repeats

Alignment Scores:

Pred. No.:	3,27e-13	Length:	1549
Score:	380.50	Matches:	133
Percent Similarity:	43.59%	Conservative:	129
Best Local Similarity:	22.13%	Mismatches:	194
Query Match:	15.31%	Indels:	145
DB:	1	Gaps:	18

US-09-837-344-41 (1-1482) x A40691 (1-1549)

```
OY      1  CAAGAACAAACAAACGATCTAGAACAGAGACGCTCTAAGAAAAGTTGCAAGAACAA  60
      ::  |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     197  GUAAGGUAAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG  216
OY      61  CAAGCGATTAA-----GAACAAGATGACTTCTAAGAAAAGTTACAA  105
      ::  |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     217  GINGUGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG  236
OY     106  GAGCGCAAGACGATTTAAGAACAGAGACGCTTGAAGAAAAGTTGCAAGAACAA  165
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     237  GUGUGUGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG  256
OY     166  ACCGATCTAAGAACAGAGACGCTGCTAAGAAAAGTTGCAAGAACAAACGATTTA  225
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     257  AAGAGUGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG  276
OY     226  GAACAAGAGAGACGCTCTAAGAAAAGTTGCAAGAACAA-----  264
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     277  GUGUGUGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG  296
OY     265  ---CAAGCGATTTAAGAACAGATGACTTCTAAGAAAAG-----  303
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     297  CysGlnTrpGlnLeuSerGlnValaAspAlaArgGlnArgValaTyrSerArgPro  316
OY     304  ---TTACAGAGACGCAAGAC-----GATTAGAACAAAGAG  336
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     317  HsArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  336
OY     337  AAGCGTCTAAGAAAAGTTGCAAGAACAAACGCGAT-----TTAGAA  381
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     337  GINGlnIleSerGlnGlnValaGlnSerLeuGlnGlnIleArgGlnArgLeuLys  356
OY     382  CAAGAGAGACGT-----  393
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     357  GINGUGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG  376
OY     394  -----GCTAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAG  438
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     377  ArgTyrThrLeuTyrAlaLysProAlaGlnArgGlnGlnValaArgGlnGlnGlnGln  396
OY     439  AAGCTTGTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAGAGAGAG  492
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     397  ArgLeuLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  416
```

```
OY     493  -----CGTCTAAGAAAAGTTGCAAGAACAAACGCGATTTAGAA  534
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     417  ArgGlnValGlnLeuGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  436
OY     535  CAAGAGACGCTCTAAGAAAAGTTGCAAGAA-----  567
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     437  ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  456
OY     568  -----CAACAAGCGATTTAGAACAGAGACGCTGCTAAGAAAAGTTGCAAGAG  618
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     457  GlnLeuTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  476
OY     619  CAGCAAGAGATTTAGAA-----CAAGAAGAGCT  648
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     477  GlnLysGlnTyrLeuGlnLysValaGlnLeuArgGlnGlnGlnGlnGlnGlnGlnGln  496
OY     649  GATACGAAAAAAATTTAGAAAAGAAAAG-----GAACATGAGATATATTAGCA  699
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     497  ArgGlnLysArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  516
OY     700  GAGCATTTATATGCGCTTGAGAAATACACGCTATAGAACTTCATCAGAAAATGACGT  759
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     517  GINGUGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG  534
OY     760  GGAATATTAT-----ATACCATCATCAATCTTTTACCTCAGAGAACAGAGCG  807
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     535  GlnTyrLeuGlnLysValaGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  554
OY     808  AATAGTAGAGATTCAGAGAAAATATCTAATAGAAAAAAACAATAGAGAA-----  858
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     555  LysArgArgGlnGlnLysGlnLysGlnTyrLeuGlnGlnGlnGlnGlnGlnGlnGln  574
OY     859  -----TCTATTACAAACAATGTTGAGAGAGAGAGATATCATTAAGACATCTGAA  912
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     575  GlnLeuGlnArgGlnGlnArgGlnLysArgArgGlnGlnGlnGlnGlnGlnGlnGln  594
OY     913  GAA-----AAGAAAGATGCTTCAATTAACCAAGAACAAAGAAATATCTGCT  963
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     595  LysValaGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  614
OY     964  GACATCAAAATATCATATCATATGACAGACAGTAAATATTTCTGATGTAATGATTTCA  1020
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     615  GlnArgGlnLysArgGlnTyrLeuGlnLysValaGlnLeuGlnGlnGlnGlnGlnGln  634
OY     1021  -----ATTAAGT  1026
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     635  GlnGlnArgGlnLysArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  654
OY     1027  AAGTATGAGGATGAATAATAGCTGTAATATGACATTCATTAATAGATGAAGAAAGAT  1086
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     655  ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  672
OY     1087  GATGAAGACTTAGACGAAATTTAAGCTTATGTGCAATATGACAAATTTCCAGATGAAGAA  1146
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     673  LysArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  692
OY     1147  AACATA-----GGAATTTATTAAGAA  1167
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     693  GlnLeuGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  712
OY     1168  CTAGAAAGATTTGATAGAGAAAATAGAAAATTTAGATGATTTAGATGAAGAAATAGAAAA  1227
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     713  LysGlnGlnLeuGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  732
OY     1228  TCATCAGAGAA-----TTATCTGAAGAAAAAATA  1257
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     733  GlnAspArgGlnLysArgArgGlnValaArgGlnArgLysTyrLeuGlnGlnGlnGln  752
OY     1258  AAAAAAGAAAAGAAATATGAAAAAAACAAGATTAATTTTAAACCAAAATGATAAAT  1317
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db     753  GlnGlnGlnLysAspArgLeuGlnArgGlnLysGlnLeuLeuArgGlnAspArgGlnLys  772
```


Oy 1318 TTGATGATGAGCATATTAAAAAATATATAAGCAAGCGTTAATACGAAAGAA 1377
 |||||
 Db 773 GlnTYrLeuGlnLysValGlnLeuGlnArgGlnGlnLeuGlnArgGlnLysArg 792
 |||||
 Oy 1378 AAA 1380
 |||||
 Db 793 Arg 793
 |||||
 RESULT 4
 A45973
 trichohyalin - human
 C:Species: Homo sapiens (man)
 C:Date: 03-May-1994 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
 A:Accession: A45973
 R:Lee, S.C.; Kim, I.G.; Marekov, L.N.; O'Keefe, E.J.; Parry, D.A.D.; Steinert, P.M.
 J. Biol. Chem. 268, 12164-12176, 1993
 A:Title: The structure of human trichohyalin. Potential multiple roles as a functional E
 ed (cross-linking) protein.
 A:Reference numbers: A45973; MUID:93280194; PMID:7685034
 A:Accession: A45973
 A:Molecule type: DNA
 A:Residues: 1-1898 <LEB>
 A:Cross-references: UNIPROT:007283; UNIPARC:UPI0000137388; GB:L09190; NID:G292835; PIDN
 C:Comment: trichohyalin is a protein of the medulla of the hair and of the inner root sh
 Covalent modifications to this protein include conversion of arginine to citrulline and
 C:Genetics:
 A:Gene: GDB:THH
 A:Cross-references: GDB:136223; OMIM:190370
 A:Map position: 1q21-q21
 C:Superfamily: trichohyalin; calmodulin repeat homology
 C:Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat
 F:49-81/Domain: calmodulin repeat homology <EF2>
 Alignment Scores:
 Pred. No.: 5 43e-13 Length: 1898
 Score: 376.00 Matches: 137
 Percent Similarity: 47.39% Conservative: 117
 Best Local Similarity: 25.56% Mismatch: 174
 Query Match: 15.12% Indels: 108
 DB: 1 Gaps: 24
 US-09-837-344-41 (1-1482) x A45973 (1-1898)
 Oy 1 CAAGAACAAACAAGCGATCTAGAACCAAGAGACGTCCTAAAGAAAGTTGCAAGACAA 60
 ::|||
 Db 522 ArgIuGlnGlnLeuArgIuArgGlnGlnGlnIuArgIuArgGlnIuArgIuArgGln 541
 |||||
 Oy 61 CAAGCGATTTAGAACCAAGATGACTCTTAAAGAAAGTTTCAAGACGACCAAGCGAT 120
 |||||
 Db 542 Gln-----GlnGlnGlnArgLeuGlnGlnArgLeuArgserGlnGlnLeuArg 558
 |||||
 Oy 121 TTAGAACAA---GAGAGACTTGCTTAAAGAAAGTTGCAAGAACCAAGCGATCTAGAA 177
 |||||
 Db 559 ArgIuGlnGlnGlnArgLeuGlnGlnLeuLeuLysArgGlnGlnLysArgLeuGln 578
 |||||
 Oy 178 CAAGAGACGTCGTAAAGAAAGTTGCAAGAACCAAGCGATCTTAAAGAAAGTTGCAAG 219
 |||||
 Db 579 GlnGlnArgIuArgGlnGlnArgLeuLysArgGlnGlnGlnGlnGlnArgIuArgGln 598
 |||||
 Oy 220 GATTTAGAACCAAGAGACGTCGTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAA 279
 |||||
 Db 599 LysArgIuGlnGlnGlnArgGlnGlnArgLeuLysArgIuGlnGlnGlnGlnArg 618
 |||||
 Oy 280 CAAGGTAGACTTGCTTAAAGAAAGTTTCAAGAGCGACGAAACGATTTAGAACAGAGAGA 339
 |||||
 Db 619 Gln--ArgLeuLysArgGlnGlnArgIuArgIuArgIuArgIuArgIuArgIuArg 633
 |||||
 Oy 340 CGTGCTTAAGAAAGTTG---CAAGAACCAAGCGATTTA--- 378
 |||||
 Db 634 ArgArgArgGlnArgLeuLysArgGlnGlnArgGlnGlnGlnGlnGlnGlnGlnGln 653
 |||||
 Oy 379 -----GAAACAAGAGACGTCGTAAAGAAAGTTGCAAGAACCAAGCGATTTA 429

[illegible]

Db 987 GIUGlUGlLeuLeuGlyGluGluProGluLySarGArGlnGluArgGluLysLys 1006
QY 1342 TATATAATGATAGACGGTTAATAAGCAAAAGAAAAATTCATATAA 1389
Db 1007 TYARxGluGluGluGluGluGluGluGluGluGluGluGluLeuArg 1022

RESULT 5
T15598
hypothetical protein C25A11.4a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15598
R:Favell, T.
submitted to: The EMBL Data Library, October 1995
A:Description: The sequence of C. elegans cosmid C25A11.
A:Reference number: Z18375
A:Accession: T15598
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1017 <FAV>
A:Cross-references: UNIPARC:UPI000017B7BA, EMBL:U39650, NID:gi1049376, PID:gi1049380, PIDN
A:Experimental source: strain Bristol N2
C:Genetic8:
A:Gene: CESP:C25A11.4a
A:introns: 65/3; 126/3; 433/2; 494/2; 711/2; 735/1; 792/3; 833/3; 873/2

Alignment Scores:
Pred. No.: 6 73e-12 Length: 1017
Score: 356.00 Matches: 120
Percent Similarity: 47.80% Conservatve: 141
Best Local Similarity: 21.98% Mismatches: 203
Query Match: 14.32% Indels: 82
Gaps: 18

US-09-837-344-41 (1-1482) x T15598 (1-1017)

QY 4 GAACAACAAGCGATCTAGACAAGACGCTGCTAAAGAAAAGTTGCAAGAACAA 63
Db 226 GIUGlUGlUAlaGluLeuAlaArg---ArgArgAlaLeuGluLysLysArgLysAspArg 244
QY 64 AGCGATTAGAACAAAGATGACTTGCTAAAGAAAAGTTGCAAGACGCAAAAGCGATT 123
Db 245 GIUGlYAlaGluGluGluGluArgLysThrMetGluArgLeuGluArgGluLysArgLysLeu 264
QY 124 GAACAAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGAT-----CTA 174
Db 265 GIUGlArgLysLeuGluGluGluGluArgLysGluLysLeuLysLeuLysLeuLysLeu 284
QY 175 GAACAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAA 231
Db 285 GIUGlArgLysArgLysLeuGluGluGluGluGluGluGluGluGluGluGluGluGlu 304
QY 232 GAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGATGACTT 291
Db 305 GIUGlArgLysLeuGluGluGluGluArgLysGluLysLeuLysLeuLysLeuLysLeu 324
QY 292 GCTAAAGAAAAGTTGCAAGAGCAAGCGATTGCAAGAGAGCGTCTAAAGAA 351
Db 325 LeuArgLysLeuGluGluGluGluArg---LeuLysArgLysArgLysLeuGluGlu 342
QY 352 AAGTTCAGAACAAACAAAGCGATT-----GAAACAAGAGAGCTGCTAAAGAA 402
Db 343 AlaArgLysLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 362
QY 403 AAGTTCAGAACAAACAAAGCGATT-----GAAACAAGAGAGCTT----- 444
Db 363 ArgLeuGluLysArgLysLeuLysArgLysArgLysLeuGluGluGluGluGluGluGlu 382
QY 445 -----GCTAAAGAAAAGTTGCAAGAACAAACAAAGC 474
Db 383 LeuValAlaAspValHisArgLysLeuGluGluGluGluGluGluGluGluGluGluGlu 402
QY 475 GATTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAACAAAGCGATTAGAA 534

Db 403 ArgGluGluAlaGluArgLeuGluGluGluGluGluGluGluGluGluGluGluGluGlu 422
QY 535 CAA-----GAGAGCGTCTAAAGAAAAGTTGCAAGAACAAACAAAGCGATTAGAA 585
Db 423 ArgLysLeuAlaGluArgLysGluGluGluGluGluGluGluGluGluGluGluGluGlu 442
QY 586 -----CAAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAGCAG 621
Db 443 LeuLysLeuAlaAlaArgArgLysLysLeuLysLeuLysLeuLysLeuLysLeuLysLeu 462
QY 622 CAAAGAGTTTAAAGAACAAAGAGCGTCTGATCGAAAAAAATTTGAAAGAAAAAGCA 681
Db 463 GIUGlArgLys---ArgValArgGluGluGluGluGluGluGluGluGluGluGluGluGlu 481
QY 682 CATGAGATATATTAAGCAGAGATTTATATAGTGGCTTTAGAAATTCAGGCTATAGACT 741
Db 482 GIUGlArgArgLysAlaAlaGluGluGluGluGluGluGluGluGluGluGluGluGlu 501
QY 742 CCATCAGAAAAATGAA-----CGTGGATTTATATATCATCATCATCTTCT 786
Db 502 AlaArgLeuLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 521
QY 787 TTACCTCAGACAAACAGAGGAAATGTAAGATTCGAAAGAAATTCATTAATAGAAAA 846
Db 522 LeuAspArgGluArgGluArgLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLys 541
QY 847 ACAATATAGAAATCATATTAACAACAAATGTTGAAAGCAGAGGATATCATATAGAGCAT 906
Db 542 GIUGlMetArgLys-----LysGluArgArgArgArgArgGluArgLysAlaAsn 555
QY 907 CTTGAA-----GAAAGAAAGATGCTTCAATTAACCA 939
Db 556 GIUGlUAlaGluLeuAlaAspLeuLeuGluGluGluGluGluGluGluGluGluGluGlu 575
QY 940 GAACAACAAAGAGATTAATCTGCTGACATACAAAATCTACATTGAGACAGTAATATT 999
Db 576 AengGluArgArgGluAlaValGluArgAlaAsnAsnAsnArgLeuGluLysAspArgSer 595
QY 1000 TCTGATGTTAATGATTTCAAAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1059
Db 596 ArgAspLysLeuAspHisLysLeuArgLysArgLysLeuGluGluGluGluGluGluGlu 615
QY 1060 GAT---TATTAATGATGAGAGAGAA-----GAT 1086
Db 616 LysArgArgLeuLeuAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 635
QY 1087 GATGAAGCTTAAGCAATTTAAGCTTATGTCATATGACAAATTTCCAGATGAAAGA 1146
Db 636 SerGluThrLeuAlaLysLeuThrGlnProMetGlyTyrThrArgGluProGluVal 655
QY 1147 AACATAGGAATTTAAGAACTAGAGATTTGATAGAGAAATTAAGAAATTAAGATGAT 1206
Db 656 ThrThrLysValGluArgGluAlaLysLeuGluGluGluGluGluGluGluGluGluGlu 675
QY 1207 TTAGATGAAGGAATGAAAAATCATCAGAGATTTATGTAAGAAAAATTAAGAAAGGA 1266
Db 676 ValProTyrAlaProSerGlnSerAlaMetGlyTyrLysAspAsnAsp----- 691
QY 1267 AAGAAATGAAAAAACAAGAGATTAATTTTAAACCAATGATTAAGAAATTTGATGAT 1326
Db 692 -----GluAsnAsnAspArgLysTyrAsnProAsnAspLysAsnArgLysGlu 708
QY 1327 GAGCATATTAATAAATAATAAATATTAAGCAGTTAATTAAGAAAGAAAGAAATTCATA 1386
Db 709 SerSerArgSerArgTyrGlnArgAlaLys-----AsnLysLysAlaArg 723
QY 1387 AATCATTTGTTCAATATATTT---GACGAGAGCAATGAATTTTACAGATCGTGATGAG 1443
Db 724 ArgAspPheTyrHisSerSerGlnAspSerAlaAspProValThrGluArgPheArgLys 743
QY 1444 TTATCTGAAGATATACT 1461

Db 744 SerThraSraPLeuThr 749

RESULT 6

T15597

hypothetical protein C25A11.4b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15597

R:Ravello, T.

submitted to the EMBL Data Library, October 1995
A:Description: The sequence of C. elegans cosmid C25A11.

A:Reference number: Z18375

A:Accession: T15597

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-737 <FAV>

A:Cross-references: UNIPARC:UPI0001787B; EMBL:U39650; NID:g1049376; PID:g1049381; PIDN

C:GeneticB:

A:Gene: C25A11.4b

A:Introns: 65/3; 126/3; 433/2; 494/2; 711/2

Alignment Scores:

Pred. No.:	9-66-12	Length:	737
Score:	353.50	Matches:	117
Percent Similarity:	47.81%	Conservative:	134
Best Local Similarity:	22.29%	Mismatches:	193
Query Match:	14.22%	Indels:	81
DB:	2	Gaps:	17

US-09-837-344-41 (1-1482) x T15597 (1-737)

```
Oy 4 GAACACAAAGCGATCTTAAGAAACAAGAGACGCTGTAAGAAAGTTGCAACAAACA 63
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 226 GUGUGUGUUAAGUUAAGUUAAGUUAAGUUAAGUUAAGUUAAGUUAAGUUAAGUUAAG 244
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 64 AGCGATTAGAAACAAGATGACCTGCTTAAGAAAGTTGCAACAAAGCGATTTA 123
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 245 GUGUUAAGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 264
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 124 GAACAGAGAGCGCTGTAAGAAAGTTGCAACAAACAAGCGAT-----CTA 174
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 265 GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 284
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 175 GAACAGAGAGCGCTGTAAGAAAGTTGCAACAAAGCGATTTAAGAAACA 231
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 285 GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 304
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 232 GAGAGACGCTGTAAGAAAGTTGCAACAAAGCGATTTAAGAAAGTTGACTT 291
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 305 GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 324
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 292 GCTAAAGAAAGTTGTAAGAGCAAGCGATTTAAGAAAGAGAGCGCTGTAAGAA 351
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 325 LEUUAAGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 342
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 352 AAGTTGCAACAAACAAGCGATTTA-----GAACAAAGAGAGCGCTGTAAGAA 402
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 343 AUAAGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 362
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 403 AAGTTGCAACAAACAAGCGATTTA-----GAACAAAGAGAGCTT----- 444
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 363 ARGLEUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 382
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 445 -----GCTAAAGAAAGTTGCAACAAACAAGCGAT 474
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 383 LEUUAUAUAAGUUAUAAGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 402
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 475 GATTAGAAACAAGAGCGCTGTAAGAAAGTTGCAACAAACAAGCGATTTAGAA 534
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 403 ARGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 422
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 535 CAA-----GAGAGAGCGCTGTAAGAAAGTTGCAACAAACAAGCGATTTAGAA 585
```

```
Db 423 ARGILEAPALAGUUAAGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 442
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 586 -----CAAGAGAGCGCTGTAAGAAAGTTGCAACAGACAG 621
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 443 LEUUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 462
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 622 CAAAGAGATTTAGAAACAAGAGCGCTGTAAGCAAGAAAGTTGAGAAAGAAAGAAAGAA 681
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 463 GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 481
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 682 CATGAGATATTAATTAAGAGAGATTTATATGCTGTTGAAATTAACCGATTAAGAACTT 741
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 482 GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 501
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 742 CCATCAGAAATGAA-----CGTGAATATTAATACCAATCAATCTTCT 786
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 502 AUAAGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 521
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 787 TTACCTCAGACAAACAAGAGATTTAGATTTCCAAAGAAATATCTATTAAGAAAGAA 846
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 522 LEUUAAPARGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 541
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 847 ACNAATGAGAGATCTATTAACAAGATTTGAAGAGCAAGAGCGATTAAGAAAGCAAT 906
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 542 GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 555
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 907 CTGGA-----GAAAGAAAGATGCTTAATTAAGCA 939
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 556 GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 575
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 940 GAACAAAGAAAGATTAATCTGCTGACATCAATCAATCAATCAATCAATCAATCAAT 999
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 576 ASNGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 595
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 1000 TCGATGCTTAATGATTTTCAATTAAGATGAGATGAGAGAAATTAAGCTGCAATTAAGAC 1059
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 596 ARGAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRY 615
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 1060 GAT-----TCATTAATTAAGTAAGAAAGAA-----GAT 1086
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 616 LYSARGARGLEUUAAGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 635
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 1087 GATGAAGCTTAAGCAAGATTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCAAG 1146
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 636 SERGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 655
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 1147 AACATAGGAATTTAAGAAAGCTAAGAGATTTGATGAGAAAGAAAGTTGATGATGAT 1206
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 656 THRTRUYVALGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 675
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 1207 TTGATGAGAGATTAAGAAAGATCTCAAGAAATTAATCGAAGAAAGAAAGAAAGAA 1266
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 676 VALPROTYRALPROSERGINSERVALMECGLTYRLEUAPRYLEUAPRYLEUAPRY 691
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 1267 AAGAAATATGAAAGAAAGAGATTAATTTAAACCAATGATTAAGTTGATGAT 1326
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 692 -----GUAENAPARGAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRY 708
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 1327 GAGCATATTAAGAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1386
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 709 SERSERARGSERARGTYRGLAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRY 723
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Oy 1387 AATATCATTTGTTTCAT 1401
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db 724 ARGAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRYLEUAPRY 728
    |||.....|||.....|||.....|||.....|||.....|||.....|||.....
```

RESULT 7

A71623

secreted protein PB0115w - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004

C:Accession: A71623


```

Db      ||||| ||||| |||||
348 AlaGluValLeuThrGlyThrTyrGluSerHisValLeuGluThrTyrLeuLeuSer 367
Qy      724 -----ATACCACTATAGAACTTCCATGAGAAATGAAAGCT 759
Db      368 GluMetThrArgMetAenAenSerLeuThrGlnTyrLeuGluSerGluSer 387
Qy      760 GGATATTATAT-----CCACATCAATCTTTTAACTCCAGACACAGAGGG 807
Db      388 GLuTyrPheValArgGlnLeuGluGluTyrSerLeuLeuSerGlnLeuSerArgGly 407
Qy      808 AATGTAGAGATCCAGAGAAAT----- 831
Db      408 LysThrSerPheThrGlnGlnIleGluGluLeuArgGlnLeuGluGluThrLys 427
Qy      832 -----TCTATATATGAA 843
Db      428 SerLysAenAlaLeuAlaLeuAlaLeuGlnAlaAlaArgHisValArgSerLeuLeuArg 447
Qy      844 AAAACCAATAGAGAAATCTATTATCAACAAATGTTGAGGACGAGGATATCATTAAGGA 903
Db      448 GLuGlnTyrGluGluGluGlnGluGlnAlaLysValaGlnLeuGlnArgAlaLeuSerLysGly 467
Qy      904 CATCTGTAA-----GAAAAGAAAGATGGTTCAATAAACCAAGAA 942
Db      468 AenAlaGluValAlaGlnTyrArgThrTyrGluTyrPheArgAlaIleGlnArgThrGln 487
Qy      943 CAAAAGAGATAAATCTGCTGACATA-----CAAAATCATACATTAGACAGACTA 993
Db      488 GLuLeuGluAenArgAlaLysLysLysLeuLeuAlaArgLeuGlnGlnAlaGluGlnAlaIle 507
Qy      994 AATATTTCTGATGTTAATGATTTTCAATAGTAAGTATGAGATGAAATATAGCTCGAA 1053
Db      508 GLuAlaAlaAenAlaLysCysSerSerLeuGluLysAlaLysHisValArgLeuGlnAenGln 527
Qy      1054 TATGAGCTTCATTAATGATGAGAGAGATGATGAGAGCTTAAGAGAAATTTAAGCTT 1113
Db      528 GlnGluAenPheMetIleAenSerLeuLysAlaAenSerAlaAlaAenSerLeu----- 545
Qy      1114 ATTGTGCATATGACAAATTTCCAGATGAGAGAAACATAGGAATTTATTAAGAACTAGAA 1173
Db      546 -----AenLysLysGlnArgGlnPheAenLysIleIleLeuSer 557
Qy      1174 GATTGTAGAGAAATGAAATTTAGATGATTAAGTGAAGCAATGAAATATCATCA 1233
Db      558 AenTyrLysGlnTyrGluGluSerGlnAlaGluLeuGlnAlaSerGlnLysGlnAla 577
Qy      1234 GAAGAAATTTATGAGAAATTAATAAAGAAAGAAATTTGAAAAACAAAGATATAT 1293
Db      578 ArgSerLeuSerThrGluLeuPheLysLeuLysAenAlaTyrGluGluThrLeuAenPheHis 597
Qy      1294 -----AATTTAAACCAATGATTAAGTTGTATGATGAG----- 1329
Db      598 LeuGluThrLeuLysValArgGluAenLysAenLeuGlnGluIleSerAenPheLeuThrSer 617
Qy      1330 -----CATATTAATAAATATTAATAATGATTAAGCAG 1359
Db      618 GlnIleSerGluGluAenLysAenLeuIleGluIleGluLysValLys-----LysGln 635
Qy      1360 GTTAATAAGAAAGAAAGAA 1380
Db      636 ValGluGlnGluLysSerGln 642

```

```

A:Reference number: A33430; MUID:90026426; PMID:2803315
A:Accession: A33430
A:Molecule type: mRNA
A:Residues: 1-771 <HA>
A:Cross-references: UNIPROT:P12957; UNIPARC:UPI0000126E45; GB:M28417; NID:G211895; PIDN
A:Experimental source: gizzard
A>Note: part of this sequence was confirmed by protein sequencing
R:Bryan, J.; Imai, M.; Lee, R.; Moore, P.; Cook, R.G.; Lin, W.G.
J. Biol. Chem. 264, 13873-13879, 1989
A:Title: Cloning and expression of a smooth muscle caldesmon.
A:Reference number: A32642; MUID:89340480; PMID:2760048
A:Accession: A32642
A:Molecule type: mRNA
A:Residues: 1-318,334-771 <BRY>
A:Cross-references: UNIPARC:UPI000001713AB; GB:U04968; NID:G212656; PIDN:AAA49067.1; PID
A>Note: the authors translated the codon GAA for residue 743 as Leu
R:Hayashi, K.; Yamada, S.; Kanda, K.; Kimizuka, F.; Kato, I.; Sobue, K.
Biochem. Biophys. Res. Commun. 161, 38-45, 1989
A:Title: 35kDa fragment of h-caldesmon conserves two consensus sequences of the tropomy
A:Reference number: A32445; MUID:89273666; PMID:2730665
A:Accession: A32445
A:Molecule type: mRNA
A:Residues: 466-771 <HA2>
A:Cross-references: UNIPARC:UPI000015624D; GB:M26684; NID:G211897; PIDN:AAA48811.1; PID
R:Maek, A.S.; Carpenter, M.; Smillie, L.B.; Wang, J.H.
J. Biol. Chem. 266, 19971-19975, 1991
A:Title: Phosphorylation of caldesmon by p34(cdc2) kinase. Identification of phosphoryl
A:Reference number: A41064; MUID:92041815; PMID:1939059
A:Accession: A41064
A:Molecule type: protein
A:Residues: 597-600,678-696,711-721 <MAK>
A:Cross-references: UNIPARC:UPI0000173E63; UNIPARC:UPI0000173E64; UNIPARC:UPI0000173E65
R:Wang, C.L.A.; Wang, L.W.C.; Lu, R.C.
Biochem. Biophys. Res. Commun. 162, 746-752, 1989
A:Title: Caldesmon has two calmodulin-binding domains.
A:Reference number: A60461; MUID:89334885; PMID:2757638
A:Accession: A60461
A:Molecule type: protein
A:Residues: 2-17, 'X', 19-38,466-485 <WAN>
A:Cross-references: UNIPARC:UPI0000173E66; UNIPARC:UPI0000173E67
R:Hatanu, M.; Hayashi, K.; Yano, H.; Takeuchi, O.; Sobue, K.
Biochem. Biophys. Res. Commun. 197, 145-153, 1993
A:Title: Common structural and expressional properties of vertebrate caldesmon genes.
A:Reference number: PC2003; MUID:94071934; PMID:8250919
A:Accession: PC2003
A:Molecule type: DNA
A:Residues: 74-419 <HAR>
A:Cross-references: UNIPARC:UPI0000173E68
R:Takeagi, T.; Yazawa, M.; Ueno, T.; Suzuki, S.; Yagi, K.
J. Biochem. 106, 778-783, 1989
A:Title: Amino acid sequence studies on cyanogen bromide peptides of chicken caldesmon
A:Reference number: PX0022; MUID:90130380; PMID:2613684
A:Accession: PX0022
A:Molecule type: protein
A:Residues: 462-477, 'D', 479-563,674-762, 'A', 763-771 <TRK>
A:Cross-references: UNIPARC:UPI0000173E69; UNIPARC:UPI0000173E6A
C:Comment: This protein plays a vital role in the regulation of smooth muscle and nonmu
C:Comment: The binding of caldesmon to F-actin is modulated by calcium and calmodulin.
C:Comment: Two calmodulin molecules can bind to nonoverlapping domains of each caldesmo
C:Superfamily: caldesmon
C:Keywords: actin binding; alternative splicing; calmodulin binding; cytoskeleton; musc
F:1-771/Product: h-caldesmon #status predicted <HMMAT>
F:1-318,334-771/Product: h-caldesmon, alternative splice form #status predicted <HMMAT>
F:266-390/Region: 13-residue repeats
F:511-582/Region: tropomyosin binding
F:622-636/Region: tropomyosin binding
F:597,662,717/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #status experim
F:688,711/Binding site: phosphate (Thr) (covalent) (by cdc2 kinase) #status experimen

```

```

Alignment Scores:
Pred. No.: 4,8e-10 Length: 771
Score: 321.00 Matches: 122

```


QY 109 CAGCAA-----AGCGAT 120
: : : : :
Db 972 GluLysValThrAlaGluAlaLysLLeLysLLeuGluLysAspIleLeuValMetAsp 991
QY 121 TTGAACAAGAGAGACTTGTCTAAAGAAAG-----TTGCAAGAACCAACAAAGCGATCTTA 174
: : : : :
Db 992 AspGlnAsnSerLysLysLeuSerLysGluArgLysLysLeuGluLysValSerAspLeu 1011
QY 175 -----GAAACAAGAGAGAGCGTCTAA----- 195
: : : : :
Db 1012 ThrThrAsnLeuAlaGluGluGluLysAlaLysAsnLeuThrLysLysLys 1031
QY 196 -----GAAAGCTTGCAGAACCAACAAAGCGATTTTGAACAAGAGAGCGTCTTAAGAA 249
: : : : :
Db 1032 HisGluSerMetLysSerGluLysLysValArgLysLysLysGluLysSerArgGln 1051
QY 250 AAGTTCAGAACCAACAAAGCGATTTTGAACAAGAT----- 285
: : : : :
Db 1052 GluLeuGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1071
QY 286 -----AGACTTGCTTAA-----GAAAG 303
: : : : :
Db 1072 AlaAspLeuGlnAlaGlnIleAlaGluLysLysLysLysLysLysLysLysLys 1091
QY 304 TTACAAAGAGAGAGAGCGATTTTGAACAAGAG----- 336
: : : : :
Db 1092 LeuGlnAlaAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1111
QY 337 -----AGACGTGCTTAAAGAAAGTTTGCAGAACCAACAAAGCGATTTTGAACAAGAGAG 390
: : : : :
Db 1112 LysIleArgGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1131
QY 391 CGTGTAAAGAAAGTTTGCAGAACCAACAAAGCGATTTTGAACAAGAGAGAGCTTCT-- 447
: : : : :
Db 1132 AlaAlaArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1151
QY 448 AAGAAAGAAAGTTTGCAGAAC-----CAACAAAGCGATTTTGAACAAGAGAGAGCGTCTAA 501
: : : : :
Db 1152 LysThrGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1171
QY 502 GAAAG-----TTGCAAGAACCAACAAAGCGATTTTGA 534
: : : : :
Db 1172 ArgGluGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1191
QY 535 -----CAAGAGAGAGCGTCTTAA-----GAAAGTTTGCAGAACCAACAAAGCGAT 579
: : : : :
Db 1192 AlaGlnValAlaGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1211
QY 580 TTGAACAAGAGAGAGCGTCTTAAAGAAAGTTTGCAGAGAGAGAGAGAGAGAGAGAGAG 639
: : : : :
Db 1212 LeuGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1231
QY 640 AGGAAGCGTCTAT-----ACGAAGAAAGAAAGTTTGA 669
: : : : :
Db 1232 GluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1251
QY 670 AGAAGAAAG 729
: : : : :
Db 1252 HisLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1264
QY 730 GCTATAGAACTTCATCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 789
: : : : :
Db 1265 -----GlnSerLysCys 1268
QY 790 CCTCAGAGAACAG 849
: : : : :
Db 1269 SerAspLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1287
QY 850 AATAG 900
: : : : :
Db 1288 GluValGlnSerValThrGlyMetLeuAsnGluLysLysLysLysLysLysLys 1303
QY 901 GGACATCTTGAAGAAAG 960

Db 1304 -----IleLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1321
QY 961 GCTGACATACAAATATCATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
: : : : :
Db 1322 GluLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1337
QY 1021 ATAAGTATAG 1080
: : : : :
Db 1338 LeuArgGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1356
QY 1081 GAAGATGATGAG 1140
: : : : :
Db 1357 GluAlaLysGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1373
QY 1141 GAAAGAAACATGAG 1200
: : : : :
Db 1374 SerAspSer-----LysLysLysLysLysLysLysLysLysLysLysLysLysLys 1390
QY 1201 GATGATTTAGATGAG 1251
: : : : :
Db 1391 GluGluGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1410
QY 1252 AAT 1290
: : : : :
Db 1411 LysAlaAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1430
QY 1291 -----AAT 1311
: : : : :
Db 1431 AspLeuValLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1450
QY 1312 AAT 1350
: : : : :
Db 1451 LysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1469
QY 1351 -----GATTAAG 1404
: : : : :
Db 1470 GluArgAspArgAlaGluAlaGluAlaArgGluLysLysLysLysLysLysLys 1488
QY 1405 TTTGACGAG 1446
: : : : :
Db 1489 AlaArgAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1502

RESULT 12
JCS420
smooth muscle myosin heavy chain 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004
C:Accession: JCS420
R:Heegawa, K.; Atakawa, E.; Oda, S.; Matsuda, Y.
Biochem. Biophys. Res. Commun. 232, 313-316, 1997
A:Title: Molecular cloning and expression of murine smooth muscle myosin heavy chains.
A:Reference number: JCS420; NUID:97242182; PMID:9125171
A:Accession: JCS420
A:Molecule type: mRNA
A:Reichers: 1-1972 <HAS>
A:Cross-references: UNIPROT:O08638; UNIPARC:UPI00000279BF; DDBJ:D85923; NID:q1945077; P
A:Experimental source: smooth muscle
C:Comment: This protein plays a role in smooth muscle cell contraction.
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: nucleotide binding; P-loop
F:88-771/Domain: myosin motor domain homology <MOT>
F:178-185/Region: nucleotide-binding motif A (P-loop)

Alignment Scores:
Pred. No.: 5,21e-10 Length: 1972
Score: 319.00 Matches: 144
Percent Similarity: 42.02% Conservative: 114
Best Local Similarity: 23.45% Mismatches: 160
Query Match: 12.83% Indels: 176
DB: 2 Gaps: 27

US-09-837-344-41 (1-1482) x JCS420 (1-1972)

[illegible]

QY 790 CCTCAGACACAGAGGAAATAGTAGAGTTCCAGAGAAATATCATATATACAAAAACA 849

Db 1269 SerAepGlyGluArgAlaArgAlaGluSerAepLysValHisLysLeuGln---Aan 1287

QY 850 AATAGAGAAATCTTTTACA-----ACAAATGTTGAGAGACGAGAGGATATACATAAA 900

Db 1288 GluValGlnSerValThrGluMetLeuAenGlnLagluGluLysValA----- 1303

QY 901 GGACATCTTGAGAAAAGAAAGATGGTTCATATAAACGAAACAAAGAGATAAATC 960

Db 1304 -----IleLysLeuAlaLysAerValAlaSerLeuGlySerGlnLeuGlnAerThrGln 1321

QY 961 GGTGACATCAAAATCATTCATTAGACAGATTAATTTTGTGATGTTAATGATTTTCAA 1020

Db 1322 GluLeuLeuGlnGluGluThrArgGlnLysLeuAenValSerThr-----Lys 1337

QY 1021 ATTAGTAAGTATGAGATGAGATGAATTAAGTGTGATGATGACGATTCATTATAGAGAA 1080

Db 1338 LeuArgGlnLeuGlnLysArgGlnLysArgAenSer---LeuGlnAerGlnLeuAerGlnLysMet 1356

QY 1081 GAAGATGATGAAGACTTAAAGCGAATTTAAGCCTATTGTGCAATATGACAAATTTCCAAAGT 1140

Db 1357 GluAlaLysGlnAenLeuGlnArgHis-----ValSerThrLeuAenIleGlnLeu 1373

QY 1141 GAAGAAACATAGGAATTTTAAAGAACTAGAAATTTGATAGAGAAAAATGAAATTTA 1200

Db 1374 SerAepSer-----LysLysLysLeuGlnAerPheAlaSerThrIleGluValMet 1390

QY 1201 GATGATTTTGAATGAGAAATGAGAAATATCATCAAGAGATTTATCT-----GAGAA 1251

Db 1391 GluGluGluLysLysAerGlnLysGlnLysGlnMetGluGlyLeuSerGlnGlnIrrGluGln 1410

QY 1252 AAAATTAAGAAAGAAATATGAAAAAACAAGAT----- 1290

Db 1411 LysAlaAlaAlaLysAerLysLysLeuGlnLysThrLysAerAerGlnGlnLysLeuAer 1430

QY 1291 -----AATTAATTTAAACCAATGAT 1311

Db 1431 AepLeuValAlaAepLysAerAenGlnArgGlnLeuValSerAenLeuGlnLysLysGln 1450

QY 1312 AAAAGTTGATGATGAGATTTTAAATAATATATAAAT----- 1350

Db 1451 LysLys---PheAerGlnLeuLeuAlaGlnLysAenIleSerSerLysIrrAlaAer 1469

QY 1351 -----GATTAAGCAGGTTTAAAGAAAGAAATTTCAATAAATCATTTGTTCCATATA 1404

Db 1470 GluArgAerArgAlaGlnLysAlaLysGlnLysGlnLysIrrLysAlaLeu---SerLeu 1488

QY 1405 TTTCACGAGACATGAATTTTACAGATCGTGATGATGATTA 1446

Db 1489 AlaArgAlaLeuGlnGlnLysAlaLysGlnLysLeu 1502

RESULT 13

A61231

Myosin heavy chain nonmuscle form A - human

Myosin alternate names: cellular myosin heavy chain; myosin type 9; MNMHC-A

N:Contains: myosin ATPase (EC 3.6.4.1)

C:Species: Homo sapiens (man)

C:Date: 12-May-1994 #sequence revision 14-Jul-1994 #text_change 31-Dec-2004

C:Accession: A61231, A34876, I52562, I61692

R:Simons, M.; Wang, M.; McBride, O.W.; Kawamoto, S.; Yamakawa, K.; Adelairein, C.

A:Title: Human nonmuscle myosin heavy chains are encoded by two genes located on different

A:Accession: A61231, MIMD:91316803; PMID:1860190

A:Accession: A61231

A:Molecule type: mRNA

A:Residues: 1-715 51M-

A:Cross-references: UNIPROT:P35579; UNIPROT:Q9UMJ0; UNIPARC:UP1000016AD3; GB:M69180; N

R:Saenz, C.G.; Myers, J.C.; Shows, T.B.; Leinwand, L.A.

P:Proc. Natl. Acad. Sci. U.S.A. 87, 1164-1168, 1990

A:Title: Human nonmuscle myosin heavy chain mRNA: generation of diversity through alteri

A:Reference number: A34876; MIMD:90138958; PMID:1967836

A:Accession: A34876

Db 1380 CysLeuGluThrAlaGluGluValIySaRgLySeuGluLysAspLeuGluGlyLeuSer 1399
||| :|||||
1246 -----GAAGAAAAATATAAAAGCAAGAAATATGAAAACCAAGCATTAAT 1296
||| ||||| :|||
Db 1400 GlnAghTseGluGluValAlaLaIaTyraPrySeuGluLysThyThrArgLeu 1419
||| ||||| :|||
Qy 1297 TTTAAACCAATGATTAAGATTGTGATGAT--GAGCATTATTAATAATTAATGAT 1353
||| ||||| :|||
Db 1420 GlnGlnGluLeuAspAspLeuValaIySaRgLeuAspHis-----GlnArgGlnSerAla 1437
::: ||||| :|||
Qy 1354 AAGCAGCTTAATAAGAAAGAAATTC 1383
::: ||||| :|||
Db 1438 CysAsnLeuGluLysGluLysLysPhe 1447
||| :|||||

RESULT 14
H86341
hypothetical protein P9H16_10 - Arabidopsis thaliana
CSpecies: Arabidopsis thaliana (mouse-ear cress)
CDate: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
CAccession: H86341
RTheologian, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,
Nure, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
AAuthors: Hunter, J.D.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Malil, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
AAuthors: Salzberg, S.L.; Schwartz, J.R.; Shum, P.; Southwick, A.M.; Sun, H.; Tallon,
Ker, M.; Wu, D.; Yi, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
Article: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
Reference number: AB6141; PMID:11130712
Accession: H86341
Status: preliminary
Molecule type: DNA
Residues: 1-1166 <STD>
Cross-references: UNIPROT:O95TP6; UNIPARC:UPI0000047C33; GB:AE005172; NID:g4836896; P
Genetics:
Map position: 1

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:
Best Local Similarity:	41.21%	22.93%
Query Match:	12.77%	

DB: 2 Gaps: 24

US-09-837-344-41 (1-1482) x H86341 (1-1166)

Qy 4 GAACAACAAGCGACTTAGAACACAGACGCCTGTTAAAGAAAGTTGCACAGACAA 63
||| :||| :|||
Db 17 GluGluGluAlaAspLeuLysSerArgAspArgSerArgLysSerArgLysSer 36
||| :||| :|||
Qy 64 AGCGAT-----TTAGAACAAGATAGACTTGCTAA----- 93
||| :||| :|||
Db 37 LysAspLysGlySerGluLysArgArgGluLysAspArgArgLysValYsser 56
||| :||| :|||
Qy 94 -----GAAAGTTCAAGACAGACAAAGCAT 120
||| :||| :|||
Db 57 SerAspSerGluAspAspTyrAspArgAspAspAspGluArgGluLysArgLysGlu 76
||| :||| :|||
Qy 121 TTAGAACAAGACAGACTTCTAAAGAA----- 147
||| :||| :|||
Db 77 LysGluArgGluArgArgArgArgLysAspArgValLysArgArgSerGluArg 96
||| :||| :|||
Qy 148 AAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAC-----AGACGTGCT 192
||| :||| :|||
Db 97 LysSerSerAspSerGluAspValGluGluGluArgLysArgLysArgVal 116
||| :||| :|||
Qy 193 AAAAGAAAGTTGCACAGACAAAGCCGATTTAGAACAAAGACAG----- 237
||| :||| :|||
Db 117 AengLysLys-----GluArgGlyHisArgGluHisGluArgAspArgGlyLysAsp 133
||| :||| :|||

[illegible]


```
OY 540 ----- 540
Db 1190 AlaMetValAlaAlaLeuArgLysLysHISAlaAspSerMetAlaGluLeuGluGluGln 1209
OY 541 -----AGACGCTGCTAAAGAAAAAGTTGCAAGAACAAACAAAGCGATTAGAACAA 588
Db 1210 IleAspAsnLeuGlnArgValLysGlnLysLeuGluLysGluLysSerGluLeuLysMet 1229
OY 589 GAG-----AGACGCTGCTAAAGAAAAAGTTGCAAGAACAA 615
Db 1230 GluThrAspAspLeuSerSerAsnAlaGluAlaIleSerLysAlaLysGluYAsnLeuGlu 1249
OY 616 GAGCAGCAAGAGATTAGAACAAAGAAAGCGCTGATACGAAAAAAATTTGAAAGAAAA 675
Db 1250 LysMetCysArgSerLeuGlu-----AspGlnValSerGluLeuLysThrLys 1265
OY 676 AAGAACATGGAGATATATTACGACAGATTAT-----TATGCTGTTAGAAAAATA 726
Db 1286 GluGluGluGluGlnGlnArgLeuIleAsnAspLeuThrAlaGlnArgAlaArgLeuGlnThr 1285
OY 727 CCAAGCTATAGAACTTCATCAGAA---AATGAACGTGATATATATATACCATCATCT 783
Db 1286 GluAlaGluGluTyrSerArgGlnLeuAspGluLysAspAlaLeuVal-----Ser 1302
OY 784 TCTTACCTCAGACAAACAGAGAAATAGTAGATTCACAGAAATA-----TCTATA 837
Db 1303 GlnLeuSerArgSerLysGlnAlaSerThrGlnGlnIleGluGluLeuLysHISGlnLeu 1322
OY 838 ATAGAAAAAACAAAT---AGAGATCTTATTACACAAATGTTGAAAGACGAAG----- 888
Db 1323 GluGluGluThrLysAlaLysAsnAlaLeuAlaHISAlaLeuGlnSerSerArgHISAsp 1342
OY 889 ---GATATACATTAAGACATCTTGAAAGAAAAAGAAAGATGTTCAATA----- 933
Db 1343 CysAspLeuLeuArgGluGlnIntyrGluGluGluGlnGluLysAlaGluLeuGlnArg 1362
OY 933 ----- 933
Db 1363 AlaLeuSerLysAlaAsnSerGluValAlaGlnIntyrArgThrLysTyrGluThrAspAla 1382
OY 934 -----AAACCAAGAACAAAGAAAGATAAATCTGCTGACATACAAAATCAT 978
Db 1383 IleGlnArgThrGluGluLeuGluGluAlaLysLysLysLeuAlaGlnArgLeuGlnGlu 1402
OY 979 ACATTAGAGACACGTAAATTTCTGATGCTTAATGATTTTCAATAAGTAAGTATGAGCAT 1038
Db 1403 AlaGluGlnHISValaGluAlaValAsnAlaLysCysAlaSerLeuGlnLysThrLysGln 1422
OY 1039 GAAATTAAGTGTGAATATATGACGATTCATTATATAGATGAGAGAAAGAT----- 1086
Db 1423 ArgLeuGlnAsnGluValaGluLysLeuMetLeuAspValaGluArgSerAsnAlaLysCys 1442
OY 1087 -----GATGAAGCTTGAACGATTTAAGCCTATT 1116
Db 1443 AlaAlaLeuAspLysLysGlnArgAsnPheAspLysValLeuSerGluTyrLysGln--- 1461
OY 1117 GTGCATATGACAAATTTCCAAAGTGAAGAAAAACATA----- 1152
Db 1462 ---LysTyrGluGluThrGlnAlaGluLeuGluAlaSerGlnLysGlnSerArgSerLeu 1480
OY 1153 -----GGAATTTATTAAGAA-----CTAGAAAGAT 1176
Db 1481 SerThrGluLeuPheLysValLysAsnValTyrGluGluSerLeuAspGlnLeuGluThr 1500
OY 1177 TTGATAGAGAAAAATGAATAATTTA-----GATGATTTAGATGAGGAATAGAA 1224
Db 1501 LeuArgArgGluAsnLysAsnLeuGlnGlnLysLeuSerAspLeuThrGluGlnIleAla 1520
OY 1225 AAATCATCAGAAAGAAATTAATCTGAA---GAAAAAATAAAAAGAAAGAAATATGAAAA 1281
Db 1521 GluGluGluLysGlnIleHISGlnLeuGluLysLysLysGluValaGluGlnGlnLys 1540
OY 1282 ACAAAAG---GATTAATTAATTTTAAACCAATGATTAAGTTTGTATGATGAGCATATTAA 1338
```

```
Db 1541 CysGluIleGlnAlaAlaLeuGluGluAlaGluAlaSerLeuGlnHISGluGluLys 1560
OY 1339 AAAATATAAATGATTAAGCAGTTTATTAAGGAAAAAGAAATTCATAAATCATTTGTTT 1398
Db 1561 IleLeuArgIleGlnLeuGlnLysAsnGlnValLysSerGluValAspArgLys----- 1578
OY 1399 CATATATTGACGAGACAAATGAATTTTACAGATCGTGATGAGTTTATCTGAAGATATA 1458
Db 1579 ---IleAlaGluLysAspGluGluIle-----AspGlnLeuLysArgAsnHIS 1593
OY 1459 ACTAAA 1464
Db 1594 ThrArg 1595
```

Search completed: December 7, 2005, 06:57:06
Job time : 81 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 7, 2005, 06:10:32 ; Search time 199.5 Seconds
(without alignments)
10482.135 Million cell updates/sec

Title: us-09-837-344-41

Perfect score: 2486
Sequence: 1 CAAGAACACAAACGATCT.....AATATTTTATGAAACTATA 1482

Scoring table:
BLASTSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2.1/USPTO.apool.b/US09837344/runat.05122005.133406.29460/app.query.fasta_1.1671
-DB=uniprot -QMT=fastran -SUFFIX=rup -MINMATCH=0.1 -LOOPT=0 -LOOPT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=500 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09837344 @CGN 1.1 294 @runat.05122005.133406.29460 -NCPUs=6 -ICPU=3
-DEV MMAP -LARGEOUTRY -NEG_SCORES=0 -MAIT -DSPBLOCK=100 -LONGLONG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot 05.80: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2478	99.7	493 2	Q25886_PLAFA
2	2406.5	95.6	1909 2	Q25893_PLAFA
3	2375.5	95.6	1596 2	Q25893_PLAFA
4	1451	58.4	280 2	Q25844_PLAFA
5	1450	58.3	280 2	Q25844_PLAFA
6	1447	58.2	280 2	Q25847_PLAFA
7	1447	58.2	280 2	Q25848_PLAFA
8	1447	58.2	280 2	Q25849_PLAFA
9	1445	58.1	280 2	Q25851_PLAFA
10	1445	58.0	280 2	Q25851_PLAFA
11	1442	58.0	280 2	Q25850_PLAFA
12	1442	58.0	280 2	Q25846_PLAFA
13	1437	57.8	280 2	Q25853_PLAFA
14	1434	57.7	280 2	Q25854_PLAFA
15	1433	57.6	280 2	Q25845_PLAFA
16	1432	57.6	280 2	Q25900_PLAFA

17	1430	57.5	280 2	Q25855_PLAFA	Q25855 plasmodium
18	1430	57.5	280 2	Q25888_PLAFA	Q25888 plasmodium
19	1428	57.4	280 2	Q27243_PLAFA	Q27243 plasmodium
20	1426	57.4	280 2	Q25889_PLAFA	Q25889 plasmodium
21	1366	54.9	264 2	Q25887_PLAFA	Q25887 plasmodium
22	770.5	31.0	318 2	Q25887_PLAFA	Q25887 plasmodium
23	468.5	18.8	978 2	Q7RAK4_PLAFA	Q7RAK4 plasmodium
24	451	18.1	1207 2	Q4UBH4_THEAN	Q4UBH4 theileria a
25	432	17.4	905 2	Q54XN9_DICDI	Q54XN9 dicystoell
26	429.5	17.3	3427 2	Q6YK79_PLAFA	Q6YK79 plasmodium
27	428	17.2	1082 2	Q54UG8_DICDI	Q54UG8 dicystoell
28	422	17.0	2205 2	Q54C75_DICDI	Q54C75 entamoeba h
29	419	16.9	762 2	Q50RA3_ENTHI	Q50RA3 entamoeba h
30	419	16.9	1061 2	Q54WJ1_DICDI	Q54WJ1 dicystoell
31	418	16.8	1575 2	Q519G4_ENTHI	Q519G4 entamoeba h
32	411	16.5	1738 2	Q51ED7_ENTHI	Q51ED7 entamoeba h
33	403	16.2	1119 2	Q54T78_DICDI	Q54T78 dicystoell
34	398.5	16.0	1108 2	Q9ND10_BABBI	Q9ND10 babesia big
35	393.5	15.8	1508 2	Q75JP5_DICDI	Q75JP5 dicystoell
36	391.5	15.7	2349 2	Q81455_PLAFA	Q81455 plasmodium
37	391	15.7	1750 2	Q8XDH2_NEUCR	Q8XDH2 neurospora
38	389	15.6	826 2	Q9VPS3_DROME	Q9VPS3 drosophila
39	383.5	15.4	1777 2	Q54FB8_DICDI	Q54FB8 dicystoell
40	381.5	15.3	1003 2	Q91LX9_HHV8	Q91LX9 human herpe
41	381.5	15.3	1129 2	Q9OR71_HHV8	Q9OR71 human herpe
42	380.5	15.3	1407 1	TRHY_SHERP	P37709 oryctolagus
43	380.5	15.3	1549 1	Q40947_HHV8	P22793 ovis aries
44	377.5	15.2	1089 2	Q40947_HHV8	Q40947 human herpe
45	377	15.2	1036 2	Q9DUM3_HHV8	Q9DUM3 human herpe

ALIGNMENTS

RESULT 1
ID Q25886_PLAFA PRELIMINARY; PRT; 493 AA.
AC Q25886
CD 025886
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2002 (TEMBLrel. 22, Last annotation update)
DE Liver stage antigen-1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OC NCB1_taxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94267224; Pubmed=7515922;
RA Fidock D.A., Gras-Masse H., Lepers J., Brahimi K., Benmohamed L.,
RA Mellouk S., Guerin-Marchand C., Londono A., Raharimalala L.,
RA Meis J.F., Langley G., Rousillon C., Tatar A., Druille P.,
RT "Plasmodium falciparum liver stage antigen-1 is well conserved and
RT contains potent B and T cell determinants.";
RL J. Immunol. 153:190-204(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Fidock D.A., Gras-Masse H., Lepers J., Brahimi K., Benmohamed L.,
RA Mellouk S., Guerin-Marchand C., Londono A., Raharimalala L.,
RA Meis J.F., Langley G., Rousillon C., Tatar A., Druille P.,
RT "The Plasmodium falciparum liver stage antigen LSA-1 is well conserved
RT and harbors major B- and T-cell epitopes.";
RL J. Immunol. 0:0-0(1994).
DR EMBL; Z30320; CAA82975.1; -; Genomic_DNA.
FT NOV TER 1 1
SQ SEQUENCE 493 AA; 59062 MW; FE252CC44FA6996E CRC64;
Alignment Scores:
Pred. No.: 6.07e-108
Score: 2478.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 99.68%
Matches: 493
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

Db 1406 LeuGluGlnGluAArgAlaLysGluLysLeuGlnGlnGlnGlnSerAAspLeuGluGln 1425
QY 130 GAGAGACTTGTCTAAAGAAAAGTTGCAAGAACAAACGATCTTGAACAAGAGAGCGT 189
Db 1426 GluAArgAlaLysGluLysLeuGlnGlnGlnGlnSerAAspLeuGlnGlnGlnAArg 1445
QY 190 GCTTAAAGAAAAGTTGCAAGAACAAACGATCTTGAACAAGAGAGCGTCTTAAAGAA 249
Db 1446 AlaLysGluLysLeuGlnGlnGlnGlnSerAAspLeuGlnGlnGlnGlnAArgAlaLysGlu 1465
QY 250 AAGTTGCAAGAACAAACGATCTTGAACAAGAGAGCGTCTTAAAGAAAAGTTTCAAA 309
Db 1466 LysLeuGlnGlnGlnGlnSerAAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1485
QY 310 GAGCAGCAAGCGATTTAGAACAAAGAGAGCGTCTTAAAGAAAAGTTGCAAGAACAA 369
Db 1486 GluGlnGlnGlnAArgAAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1505
QY 370 AGCGATTTAGAACAAAGAGAGCGTCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 429
Db 1506 SerAAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1525
QY 430 GAAACAAGAGAGCTTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAGACAG 489
Db 1526 GluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1545
QY 490 AGACGTCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTCT 549
Db 1546 ArgAArgAlaLysGluLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1565
QY 550 AAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGT----- 597
Db 1566 LysGluLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1585
QY 597 ----- 597
Db 1586 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1605
QY 598 -----GCTTAAAGAAAAGTTGCAAGAACAGAGCAAGAA 627
Db 1606 GlnGlnAArgAAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1625
QY 628 GATTTAGAACAAAGAGCGTGTATCGAAAAAAATTTGAAAGAAAAGAAACATGGA 687
Db 1626 AspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1645
QY 688 GATTATTTAGAGAGATTTATATGTCGCTTGAAGAAATACAGGCTATGAACTTCATCA 747
Db 1646 AspAlaLeuAlaGlnAAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1665
QY 748 GAAATGAAACGTGATATTTATATATCAATCAATCTTTTACCTGACGACAAAGAGG 807
Db 1666 GluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1685
QY 808 AATAGTAGAGATTTCCAAAGAAATCTTAATAGAAAAAAATAGAAATCTTATTAACA 867
Db 1686 AAspSerAAspSerLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1705
QY 868 ACAATGTTGAAGAGCAAGAGGATATACATAAGAGACCTTTGAAGAAAAGAAAGAGT 927
Db 1706 ThrAAspAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1725
QY 928 TCAATTAAGAACAGAAACAAAGAGATTAATCTGCTGACATACAAATCATATACATTAAG 987
Db 1726 SerLysAAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1745
QY 988 ACAAGTAATATTTCTGATGTTAATGATTTTCAATAGTAAGTAAGAGATGAATAAGT 1047
Db 1746 ThrValAAspLysAAspValAAspAAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1765
QY 1048 GCTAATATGACGATTCATTAATAGATGAAGAGAGATGAAGACTTGAAGCAATTT 1107

Db 1766 AlaGluLysAAspSerLeuLysAAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1785
QY 1108 AACCTTATGTCGAAATATGCAATTTTCCAAAGAGAAACATAGGAATTTTAAAGAA 1167
Db 1786 LysAAspAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1805
QY 1168 CTAGAAGATTTGATAGAGAAAATGAAAATTTTATGATGATTTAGATGAAGAAATAGAAAA 1227
Db 1806 LeuGlnAAspLeuLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1825
QY 1228 TCATCAGAAAGATTTATCTGAAGAAAATTAAGAAAAGAAATGAAATGAAAAACAAAG 1287
Db 1826 SerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1845
QY 1288 GATTAATATTTTAAACCAATGATTAAGATTTGATGAGCATTTTAAATTTAAATTA 1347
Db 1846 AspAAspAAspAAspAAspAAspAAspAAspAAspAAspAAspAAspAAspAAspAAsp 1865
QY 1348 AATGATTAAGCAGGTTAATAGAAAAGAAAATTCATTAATCATTTGTTTCAATATTT 1407
Db 1866 AAspAAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1885
QY 1408 GACGAGACAAATGAAATTTTACAGATCGTGAATGATTAATCTGAAGATATTAATAT 1467
Db 1886 AspGlnAAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1905
QY 1468 TTTATGAACCTA 1479
Db 1906 PheMetLysLeu 1909

RESULT 3
ID Q8IJ44_PLAF7 PRELIMINARY; PRT; 1596 AA.
AC Q8IJ44;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Liver stage antigen, putative.
GN ORFNames=PF10_0356;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=2225705; PubMed=1236864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Perera M., Allen J., Selengut J., Haft D., Mather M.W., Valiya A.B.,
RA Martin D.M.A., Fairhead A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrall B.G.,
RT "Genome sequence of the human malaria parasite Plasmodium
falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AB014834; AAN35553.1; -; Genomic DNA.
SQ SEQUENCE 1596 AA; 191691 MW; C0B3A2C0F5D60AC CRC64;

Alignment Scores:
Pred. No.: 2,82e-103 Length: 1596
Score: 2375.50 Matches: 478
Percent Similarity: 94.71% Conservative: 5
Best Local Similarity: 93.73% Mismatches: 10
Query Match: 95.56% Indels: 17
DB: 2 Gaps: 1

US-09-837-344-41 (1-1482) x Q8IJ44_PLAF7 (1-1536)
QY 1 CAAGAACAAACGATCTTGAACAAGAGAGAGCGTCTTAAAGAAAAGTTGCAAGAACAA 60
Db 1087 Gln 1106

[illegible][illegible]

Accession	Gene	Protein	Score	Percent Similarity	Beat Local Similarity	Query Match	DB
OS	Plasmodium falciparum.						
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.						
OX	NCBI_TaxID=5833;						
RN	NUCLEOTIDE SEQUENCE.						
RP	MEDLINE=96065765; PubMed=7477175; DOI=10.1016/0166-6851(95)00069-D;						
RX	Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Porco M.M.,						
RA	Hawley W.A., Collins W.E., Lal A.A.;						
RT	"Sequence variations in the non-repetitive regions of the liver stage-						
RT	specific antigen-1 (ISA-1) of Plasmodium falciparum from field						
RT	isolates.";						
RL	Mol. Biochem. Parasitol. 71:291-294 (1995).						
DR	EMBL, L40888; AAC1598.1; -; Genomic_DNA.						
FT	NON_TER						
SO	SEQUENCE						
Alignment Scores:							
Pred. No.:	7,1e-60	Length:	280				
Score:	1447.00	Matches:	278				
Percent Similarity:	100.00%	Conservative:	2				
Beat Local Similarity:	99.29%	Mismatches:	0				
Query Match:	58.21%	Indels:	0				
DB:	2	Gaps:	0				
US-09-837-344-41 (1-1482) x Q25847_PLAFA (1-280)							
QY	640 AGAAGGCGTGATACGAAAAAATTGGAAGAAAAAGAAACATGAGATATTATTAGCA	699					
DB	1 ArglybAlaAspThrIlybAsnLeuGIuArglybLyGIuHhIeGIyAspValLeuAla	20					
QY	700 GAGGATTTATATGTCGTTTAGAAATACGACTATAGAACTTCCATCAGAAATGACCT	759					
DB	21 GluAspLeuIyGIyArgLeuGIuIleProAlaIleGIuLeuProSerGIuAsnGIuArg	40					
QY	760 GGAATTTATATCCCATCAATCTTCTTACCTCAGGACAGCAAGAGGAAATGTAGAGAT	819					
DB	41 GIyIyTrIyIleProHhIeGlnSerSerLeuProGIuAspAsnArgGIyAsnSerArgAsp	60					
QY	820 TCACAAGAAATCTTATATATAGAAAAACAAATAGAGATCTTATTAACAATGTTTGA	879					
DB	61 SerIyAspGIuIleSerIleIleGIuIyStrIAsnArgGIuSerIleThrIAsnValGIu	80					
QY	880 GAGCAGAGGATATACATTAAGACACATCTTGAAAGAAAGAAAGATGCTTCAATAAACCA	939					
DB	81 GIyArgArgAspIleHhIeIyAspGIuIleAsnGIuIyLeuAspGIySerIleIyAspPro	100					
QY	940 GAACAAAAAGAGATTAATCTGCTGACATACCAAAATCATACATTTAGAGACAGTAATATT	999					
DB	101 GIuGIuIyAspGIuAspIySerIleAspIleGIuIleAsnHhIeThrLeuGIuIleValAsnIle	120					
QY	1000 TCTGATGTTAAATGATTTTCAATTAAGTATAGTATGAGATGAAATAAGCTGTAATATGAC	1051					
DB	121 SerAspValAsnAspPheGIuIleSerIyTrGIuAspGIuIleSerIleGIuIyAsp	140					
QY	1060 GATTCATTATATAGATGAGAGAGAGATGATGAAGACTTAGACGAATTTAAGCCTATTGTG	1111					
DB	141 AspSerLeuIleAspGIuIleAspGIuIleAspArgGIuIleAspArgGIuIleAspArg	160					
QY	1120 CAATATGACATTTCCAAAGTGAAGAAACATAGGAATTTATTAAGAACTAGAAAGTTTG	1177					
DB	161 GIuIyTrAspAspPheGIuIleAspGIuIleAsnIleGIyIleTrIyAspGIuIleAspLeu	180					
QY	1180 ATAGAGAAAAATGAAAAATTTAGTGATTTAGATGGAAGAAATGAAAAATCATCAGAGAA	1231					
DB	181 ILeGIuIyAsnGIuIleAsnLeuAspAspLeuAspGIuIleGIuIyIleAspSerGIuIle	200					
QY	1240 TTATCTGAGAAAAATATAAAAAGGAAAGAAATATGAAAAACAAGATATATATTT	1291					
DB	201 LeuSerGIuIleGIuIleAspIyAspGIuIyIleGIuIyIleGIuIyIleGIuIyIleGIuIy	220					
QY	1300 AAACCAATGATTAAGTTTGTATGATGACATATTAAAAATATTAATATGATTAAGCAG	1351					
DB	221 IyAspProAsnAspIySerLeuIyAspGIuIleIleIyIleIyTrIyAsnAspIyGIuIle	240					

QY		1360	GTTAATTAGAGAAAGAAAAATTCATAAATCATGTTCTTACATATTTTGACGGAGACAAT	1419
Dd		241	VAlAsnlySGUlSyGSUlSyPhelIeLySeSerleuPhehiSlIePheApGIyAspSn	260
QY		1420	GAATATTTTACAGATCGGTGATGAGTTATCTCAGAATATAACTAAATATTTTATGAACTA	1479
Dd		261	GIuleIeuGInIlIevIAApGIuleSeScIdUnspRIethrIySTyrPhneMeLySeU	280
		RESULT 7		
		Q25848	PLAFA PRELIMINARY;	PRT; 280 AA.
ID		Q25848	AC	Q25848:
DT		01-NOV-1996	(TREMBLrel. 01, Created)	
DT		01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT		01-OCT-2002	(TREMBLrel. 22, Last annotation update)	
DE		Liver stage-specific antigen 1 (Fragment).		
GN		Name=LSA-1;		
OS		Plasmodium falciparum.		
OC		Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.		
OX		NCBI_TaxID=5833;		
RN		[1]		
RP		NUCLEOTIDE SEQUENCE.		
RX		MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6861(95)00069-D;		
RA		Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,		
RA		Hawley W.A., Collins W.E., Lai A.A.;		
RT		"Sequence variations in the non-repetitive regions of the liver stage-		
RT		specific antigen-1 (LSA-1) of Plasmodium falciparum from field		
RT		isolates.";		
RL		Mol. Biochem. Parasitol. 71:291-294(1995).		
DR		EMBL; L40889; AAC41599.1; -; Genomic_DNA.		
FT		NON TER		
SQ		SEQUENCE 280 AA; 32939 MW; E77080385FPAAD61 CRC64;		
		Alignment Scores:		
		Pred. No.: 7,1e-60	Length: 280	
		Score: 1447.00	Matches: 279	
		Percent Similarity: 100.00%	Conservative: 1	
		Best Local Similarity: 99.64%	Mismatch: 0	
		Query Match: 58.21%	Indels: 0	
		DB: 2	Gaps: 0	
		US-09-837-344-41 (1-1482) x Q25848_PLAFA (1-280)		
QY		640	AGGAAGCGCTGATACGAAAAAAAAATTTTGAAAAGAAAAAGAACATGAGATATATTAGCA	699
Dd		1	ArgylsaIAApRlnrILySLyASenLeGULnGrLySLySGUlHISglYAepRIleUula	20
QY		700	GAGGATTTATATGCGTCGTTTGAAATACACAGCTATAGAACTCCATCAGAAAAAGAACGT	759
Dd		21	GluspReurYrGIyArGLyeUGlIIleProAlIleGIuleuProSeGIuAsnGluArg	40
QY		760	GGATATTATATACACATCAATCTCTTTACTCGAGGACACAGAGGAAATAGTAGAT	819
Dd		41	GIyUrTyrlIePrOHISglINseSerIeuProInIsphAmBgIyAnSeSerIgrSp	60
QY		820	TCCAAGAAATATCTATATATGAAAAAACAATAGAGATCTATTAACAACAATGTTGAA	879
Dd		61	SeIrysgUlIeserIlelleGIulLyethrAsnArgIuseRIethrThAsnValGlu	80
QY		880	GGACGAAGGATATATCATTAAGAACAATCTTGAGAAAAAGAAAGATGGTTCAATAAACA	939
Dd		81	GIyArGaTgARIIleHISglYHISeIueGUlIUlySLyABpGIySerIleLySpro	100
QY		940	GAACAAAAAGAAATTAATCTGCCTGACATCAAAATCATTCATTAGAGACAGTAATATT	999
Dd		101	LYsgInUsGIuABpLySeRIaAspRIleGIImasnHsthrIeuGUlthValAsnIle	120
QY		1000	TCTGATGTTAATGATTTTCAATATAGTAAGTAGAGATGAGTAATAGTCTGAATATGAC	1059
Dd		121	SeIserpVAIsnApPhegnIIISerIySTyrlunspGIuIeserIlaGIuITyrAsp	140
QY		1060	GATTCATTATATAGTAGAAGAAAGATGATGAACTTAGACAAATTTAAAGCTATTGTG	1119

```
Db 141 AspSerLeuIleAepGluGluGluAepSpLusAepLusAepGluPheLysProIleVal 160
Qy 1120 CAATATGACAAATTTCCAGATGAAAGAAACAATGAAATTTATTAAGAACTTGAAAGATTGG 1179
Db 161 GlntYzAspAsnPhgInAepGluGluAenIleGlyIleTyrgLysGluLeuGluAepLeu 180
Qy 1180 ATGAGAGAAAAATGAAAAATTTAGATGATTGATGAGAAATGAAAAATGAAATCACTGAGAA 1239
Db 181 ILeGluLysAsnGluAenLeuAepAepLusAepGluGlyIleGluLysSerSetGluGlu 200
Qy 1240 TTAATCTGAGAAAAATTAATAAAAAAGAAAGAAATATGAAAAACAAGATTAATATTTT 1299
Db 201 LeuSerLeuGluLysIleLysLysGlyLysLysTyrgLysThrLysAspAsnAepHe 220
Qy 1300 AAACCAATGATAAAGATTGTATGATGACATATTAATAAAATATTAATGATTAACAG 1359
Db 221 LysProAenAepLysSerLeuTyrgAepGluIleIleLysLysTyrgLysAenAepLysGln 240
Qy 1360 GTTAATAAGAGAAAAATTCATTAATCATTTGTTTCATATATTGACGGAGACAAT 1419
Db 241 ValAenLysGluLysGluLysPheIleLysSerLeuPheIlePheAepGlyAspAsn 260
Qy 1420 GAAATTTTACAGATCGTGATGATGATTATCTGAAGATATACTAAATATTTTATGAAACTA 1479
Db 261 GluIleLeuGlnIleValAepGluLeuSerGluAepIleThrLysTyrgPheMetLysLeu 280
```

RESULT 8

Q25852_PLAFA
ID Q25852_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25852;

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).

GN Name=LSA-1;

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;

RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,

RA Hawley W.A., Collins W.E., Lai A.A.,

RT "Sequence variations in the non-replicative regions of the liver stage-

RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field

RT isolates."

RL Mol. Biochem. Parasitol. 71:291-294(1995).

DR EMBL; L40834; AAB59233.1; -, Genomic_DNA.

FT NON_TER 1

SQ SEQUENCE 280 AA; 32927 MW; E6648F85FBAA245 CRC64;

Alignment Scores:

Pred. No.: 7,1e-60 Length: 280
Score: 1447.00 Matches: 278
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.29% Mismatches: 0
Query Match: 58.21% Indels: 0
DB: Gaps: 2

US-09-837-344-41 (1-1482) x Q25852_PLAFA (1-280)

```
Qy 640 AGAAGGCTGATACGAAAAAATTTAGAAAGAAAAAGAAACATGAGATATATTGACA 699
Db 1 ArglybAlAePrLrLysLysAenLeuGlnLysLysLysGluIleGlyAepValIleAla 20
Qy 700 GAGATTTATATGCTGCTTTAGAAATACCACTATTAAGAACTTCATGACGAGAAAGAACT 759
Db 21 GluAepLeuTyrgLysAepLeuGluIleProAlIleGluLeuProSerGluAenGluLys 40
Qy 760 GGATATTATATACACATCAATCTTCTTACCTCAGACGACACAGAGGAATATGAGAT 819
Db 41 GlyTyrrLrIleProHLeGlnSerSerLeuProGlnAepAenAlyGlyAenSerArgAep 60
```

```
Qy 820 TCCAGGAAATATCTATTAATAGAAAAACAATAGAAATCTATTACAACAATGTTGAA 879
Db 61 SerLysGluIleSerIleIleGluLysThrAenAepGluSerIleThrAenValGlu 80
Qy 880 GAGAGAGGATATATACATTAAGACATCTTGAAAGAAAGAAAGATGGTTCAATAAAAACA 939
Db 81 GlyArgAepAepLrLrLysLysGluIleLeuGluGluLysLysAepGlySerIleLysPro 100
Qy 940 GAAACAAAAAGAAATTAATCTGCTGACATACAAAATCTATCATTAAGACACATTAATAT 999
Db 101 GluGluLysGluAepLysSerAlAepIleGlnAenIleThrLeuGluThrValAenIle 120
Qy 1000 TCGATGTTAATGATTTTCAATTAAGTAAGTATGAGATGAAATTAAGTCTGATATGAC 1059
Db 121 SerAepAlAenAepAepGlnIleSerLysTyrgLusAepGluIleSerAlAgluTyrgAep 140
Qy 1060 GATTCAATTAATGATGAGAGAAAGATGATGAAGACTTACAGCAATTTAAGCCATTATGTG 1119
Db 141 AspSerLeuIleAepGluGluGluAepAepGluAepAepGluPheLysProIleVal 160
Qy 1120 CAATATGACAAATTTCCAGATGAAAGAAACAATGAAATTTATTAAGAACTTGAAAGATTGG 1179
Db 161 GlntYzAspAsnPhgInAepGluGluAenIleGlyIleTyrgLysGluLeuGluAepLeu 180
Qy 1180 ATGAGAGAAAAATGAAAAATTTAGATGATTGATGAGAGAAATGAAAAATCATCAGAAAGAA 1239
Db 181 ILeGluLysAsnGluAenLeuAepAepLusAepGluGlyIleGluLysSerSetGluGlu 200
Qy 1240 TTAATCTGAGAAAAATTAATAAAAAAGAAAGAAATATGAAAAACAAGATTAATATTTT 1299
Db 201 LeuSerLeuGluLysIleLysLysGlyLysLysTyrgLysThrLysAspAsnAepHe 220
Qy 1300 AAACCAATGATAAAGATTGTATGATGACATATTAATAAAATATTAATGATTAACAG 1359
Db 221 LysProAenAepLysSerLeuTyrgAepGluIleIleLysLysTyrgLysAenAepLysGln 240
Qy 1360 GTTAATAAGAGAAAAATTCATTAATCATTTGTTTCATATATTGACGGAGACAAT 1419
Db 241 ValAenLysGluLysGluLysPheIleLysSerLeuPheIlePheAepGlyAspAsn 260
Qy 1420 GAAATTTTACAGATCGTGATGATGATTATCTGAAGATATACTAAATATTTTATGAAACTA 1479
Db 261 GluIleLeuGlnIleValAepGluLeuSerGluAepIleThrLysTyrgPheMetLysLeu 280
```

RESULT 9

Q25849_PLAFA
ID Q25849_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25849;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).

GN Name=LSA-1;

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;

RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,

RA Hawley W.A., Collins W.E., Lai A.A.,

RT "Sequence variations in the non-replicative regions of the liver stage-

RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field

RT isolates."

RL Mol. Biochem. Parasitol. 71:291-294(1995).

DR EMBL; L40890; AAC41600.1; -, Genomic_DNA.

FT NON_TER 1

SQ SEQUENCE 280 AA; 32966 MW; E96D255154DEA9CF CRC64;

Alignment Scores:

Pred. No.: 8,8e-60 Length: 280
Score: 1445.00 Matches: 279

Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.64% Mismatches: 0
Query Match: 58.13% Indels: 0
DB: 2 Gaps: 0

US-09-837-344-41 (1-1482) x Q25851_PLAFA (1-280)

```
QY 640 AGAAGGCTGATACGAAAAAATTATAGAAAAAGAAACATGAGATATATTACCA 699
DB 1 ArglybAlaAspThrLysAsnLeuGluArgLysGluHisGlyAspIleLeuAla 20
QY 700 GAGATTTATATGTCGTTTAAAGAAATACCACTTATAGAACTTCATTCAGAAATGAAAGT 759
DB 21 GluAspLeuTyrGlyArgLeuGluIleProAlaIleGluLeuProSerGluAsnGluArg 40
QY 760 GGAATATTATATACCATCAATCTTCTTACTCCTGAGACACAGAGGGAATGTAGAGAT 819
DB 41 GlyTyrTyrIleProHisGlnSerLeuProGlnAspAsnArgGlyAsnSerArgAsp 60
QY 820 TCCAGAAATATCTTATATAGAAAAACAAATAGAGATCTTATACCAAAATGTTGAA 879
DB 61 SerLysGluIleSerIleIleGluLysThrAsnArgLysSerIleThrThrAsnValGlu 80
QY 880 GGACGAGGATATACATTAAGACATCTTGAAGAAAAAGAAAGATGCTCAATTAACCA 939
DB 81 GlyArgGlyAspIleHisIleGlyHisIleGluGluLysLysAspGlySerIleLysPro 100
QY 940 GAACAAAAAGAAATTAATCTGCTGACATACAAATCATATACATTAAGACAGTAAATAT 999
DB 101 GluGlnLysGluAspLysSerIleAspIleGlnAsnHisThrLeuGluThrValAsnIle 120
QY 1000 TCTGATGTTATAGATTTTCAAAATAGTATAGGAGAAATAGTCTGAATATGAC 1059
DB 121 SerAspValAsnAspPheGlnIleSerLysTyrGluAspGluIleSerIleGluLysArg 140
QY 1060 GATTCATTATAGATCAAGAAAGAGATGTAAGAACTTAGACAAATTTAAGCTTATG 1119
DB 141 AspSerIleIleAspGlnGluGluAspAspGluAspGluPheLysProIleVal 160
QY 1120 CAATATGACAAATTTCCAGATGAGAAACATAGAAATTTATTAAGAACTTAGAGATTG 1179
DB 161 GlnTyrAspAsnPheGlnIleAspGlnGluAsnIleGlyIleTyrLeuGluLeuAspLeu 180
QY 1180 ATAGAGAAAAATGAAAAATTTAGATGATTTAGATGAGAAATAGAAAAATCATCAGAGAA 1239
DB 181 IleGluLysAsnGluAsnLeuAspAspLeuAspGluGlyIleGluLysSerGluGlu 200
QY 1240 TTATCTGAAGAAAAATTAAGAAAGAAATATGAAAAACAAAGATATATATTT 1299
DB 201 LeuSerGluGluLysIleLysLysGlyLysTyrGluLysThrLysAspAsnAspPhe 220
QY 1300 AAACCAATGATTAAGATTGATGATGAGCATATTAATAAATATTAATAATGATTAAGCG 1359
DB 221 LysProAsnAspLysSerLeuTyrAspGluTyrIleLysLeuTyrLysAsnAspLysGln 240
QY 1360 GTTATATAGAAAAAGAAAAATTCATTAATATCTTCTTCATATATTTGACGAGACAT 1419
DB 241 ValAsnLysGluLysGluLysPheIleLysSerLeuPheHisIlePheAspGlyAspAsn 260
QY 1420 GAAATTTTACAGATCGTGATGAGTATCTGAAGATATTAATAATTTTATGAAACTA 1479
DB 261 GluIleLeuGlnIleValAspGluLeuSerGluAspIleThrLysTyrPheLeuLysLeu 280
```

RESULT 10

Q25851_PLAFA PRELIMINARY: PRT: 280 AA.

AC Q25851_ 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
GN Liver stage-specific antigen 1 (Fragment).
OS NameLSA-1; Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6861(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins M.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates."
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL, L40893; AAC41603.1; -; Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32927 MW; 48706EFDCE40ADD3 CRC64;

Alignment Scores:
Pred. No.: 8.8e-60 Length: 280
Score: 1445.00 Matches: 278
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.29% Mismatches: 0
Query Match: 58.13% Indels: 0
DB: 2 Gaps: 0

US-09-837-344-41 (1-1482) x Q25851_PLAFA (1-280)

```
QY 640 AGAAGGCTGATACGAAAAAATTATAGAAAAAGAAACATGAGATATATTACCA 699
DB 1 ArglybAlaAspThrLysAsnLeuGluArgLysGluHisGlyAspIleLeuAla 20
QY 700 GAGATTTATATGTCGTTTAAAGAAATACCACTTATAGAACTTCATTCAGAAATGAAAGT 759
DB 21 GluAspLeuTyrGlyArgLeuGluIleProAlaIleGluLeuProSerGluAsnGluArg 40
QY 760 GGAATTTATATACCATCAATCTTCTTACTCCTGAGACACAGAGGGAATGTAGAGAT 819
DB 121 SerAspValAsnAspPheGlnIleSerLysTyrGluAspGluIleSerIleGluLysArg 60
QY 1060 GATTCATTATAGATCAAGAAAGAGATGTAAGAACTTAGACAAATTTAAGCTTATG 1119
DB 141 AspSerIleIleAspGlnGluGluAspAspGluAspGluPheLysProIleVal 160
QY 1120 CAATATGACAAATTTCCAGATGAGAAACATAGAAATTTATTAAGAACTTAGAGATTG 1179
DB 161 GlnTyrAspAsnPheGlnIleAspGlnGluAsnIleGlyIleTyrLeuGluLeuAspLeu 180
QY 1180 ATAGAGAAAAATGAAAAATTTAGATGATTTAGATGAGAAATGAAAAATCATCAGAGAA 1239
DB 181 IleGluLysAspGluAsnLeuAspAspLeuAspGluGlyIleGluLysSerGluGlu 200
QY 1240 TTATCTGAAGAAAAATTAAGAAAGAAATATGAAAAACAAAGATATATATTT 1299
DB 201 LeuSerGluGluLysIleLysLysGlyLysTyrGluLysThrLysAspAsnAspPhe 220
QY 1300 AAACCAATGATTAAGATTGATGATGAGCATATTAATAAATATTAATAATGATTAAGCG 1359
DB 221 LysProAsnAspLysSerLeuTyrAspGluTyrIleLysLeuTyrLysAsnAspLysGln 240
```

Qy	1360	CTTATATAGAGAAAAGAAAATTCCTAAATATCTGTTTCTCATATTTTGACGAGACAT	1413
Db	241	ValenSyGluYcGluYbSerHeileYbSerLeuPheHisIlePheAspGlyAspAsn	260
Qy	1420	GAATTTTACGATGCGATGAGTTATCTGAAATATACCTAAATATTTATGAACTA	1479
Db	261	GluIleLeuGlnIleValAspGluLeuSerGluAspIleTrnLYrYrPheMetYbLeu	280
RESULT 11			
Q25850_PLAFA			
ID	Q25850	PLAFA PRELIMINARY;	PRT; 280 AA.
AC	Q25850		
DT	01-NOV-1996	(TREMBlrel. 01, Created)	
DT	01-NOV-1996	(TREMBlrel. 01, Last sequence update)	
DT	01-OCT-2002	(TREMBlrel. 22, Last annotation update)	
DE	Liver stage-specific antigen 1 (Fragment).		
GN	Name=LSA-1;		
OS	Plasmodium falciparum.		
CC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI	TextID=583;	
RN	NUCLEOTIDE SEQUENCE.		
RP	[1]		
RX	MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;		
RA	Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,		
RA	Hawley M.A., Collins W.E., Lal A.A.;		
RT	"Sequence variations in the non-repetitive regions of the liver stage-		
RT	specific antigen-1 (LSA-1) of Plasmodium falciparum from field		
RL	isolates."/		
RL	Mol. Biochem. Parasitol. 71:291-294(1995).		
DR	EMBL, L40891; AAC41601.1; -, Genomic_DNA.		
FT	NON_TER		
SO	SEQUENCE	280 AA; 32989 MW; E96F812CFEABED8 CRC64;	
Alignment Scores:			
Pred. No.:	1,096-59	Length:	280
Score:	1443.00	Matches:	279
Percent Similarity:	99.64%	Conservative:	0
Best Local Similarity:	99.64%	Mismatches:	1
Query Match:	58.05%	Indels:	0
DB:	2	Gaps:	0
US-09-837-344-41 (1-1482) x Q25850_PLAFA (1-280)			
Qy	640	AGGAAGCGCTGATACGAAAAAAATTTAGAAAAGAAAAAGAACATGAGATATATAGCA	699
Db	1	ArglybAlaBpTrnLYrYbAsnLeuGluTrnGlybLYbHISgLYbAspIleLeuAla	20
Qy	700	GAGGATTTATATGCGTCTTTGAAATATCCAGCTATAGACTTCATCAGAAAATGAACT	759
Db	21	GluBpLeuTrYGlyArgLeuGluIleProAlaIleGluLeuProSerGluAsnGluArg	40
Qy	760	GGATTTATATACSCATCATCTCTTACTCTCGAGCAACAGAGGAATATAGAGAT	819
Db	41	GlyTrnLYrTrilePheHisGlnSerSerLeuProGlnBpAsnArgGlyAsnSerArgAsp	60
Qy	820	TCCAGGAATATCTATATAGAAAAAAACAATAGAGATCTATTACAACAATGTTGAA	879
Db	61	SerLYrGluIleSerIleIleGluYrThrAsnArgIuSerIleTrnThrAsnValGlu	80
Qy	880	GGACGAGGAGATATCATTAAGAGACATCTTGAGAAAAAGAAAGATGCTTCAATAAACCA	939
Db	81	GlyArgAspArgIleHisLYrGlyHisLeuGluGluYbLYbAspGlySerIleYbPro	100
Qy	940	GAAACAAAAGAGATTAATCTGCTACATCAAAATCTCATACATTAGAGACAGTAAATTT	999
Db	101	GluGlnLYrGluBpLYrSerHisAspIleGlnHisnHisTrnLeuGluTrnValAsnIle	120
Qy	1000	TCTGATGTTAATGATTTTCAATATAGTATAGATGAGATGAATATAGTCTGAATATGAC	1059
Db	121	SerArgValAsnAspPheGlnIleSerLYrYrGluBpGluIleSerAlaGluTrnAsp	140
Qy	1060	GATTCATTAAATGATGAAGAGAGATGATGAAGCTTAGACGAATTTAAAGCTATTTGTG	1119

Dd	141	Aspergillus leucospilus	160
Oy	1120	CATATGACAAATTTCCAGATGAAGAAAAACATGAGAATTATTAAGAACTAGAACATTTGG	117
Dd	161	GlnTyrAepAnPheGlnAspGluGluAsnIleGlyIleTyrLysGluLeuGluAspLeu	180
Oy	1180	ATAAGAAAATTTGAAAAATTTAGATCATTTAGATGAAGAAATTTGAAAAATTCACAGAAAGA	123
Dd	181	IleGluYrsAnGluAsnLeuAspSerLeuGluGlyIleGluLysSerSerGluGlu	200
Oy	1240	TTATCTGAAAGAAAAATATAAAAAGCAAGAAATATGAAAAACAAGATPATATTTT	129
Dd	201	LeuSerGluGluYrsIleLysLysGlyLysLysTyrGluLysThrLysAspYrAsnPro	220
Oy	1300	AAACCAAATGATAAAAAGTTTGATGATGAGCATATTAATAAAATATATAAATGATAAACAG	135
Dd	221	LysProAsnAerLysSerLeuTyrAspGluHisIleLysLysTyrLysAsnAspLysGln	240
Oy	1360	GTTAATAAGAAAAAGAAAAATTCATAAATCATTTGTTTCATATTTTGAACGAGACAT	141
Dd	241	ValAsnLysGluLysGluLysPheIleLysSerLeuPheHisIlePheAspGluAsnPro	260
Oy	1420	GAAATTTTACAGATCGTGCATGATGATTTACTGTGAAGATATTAACATTAATATTTTGAACCTA	147
Dd	261	GluIleLeuGlnIleValAspGluLeuSerGluAspIleThrLysTyrPheMetLysLeu	280
 RESULT 12 Q25846_PLAFA ID Q25846_PLAFA PRELIMINARY; PRT; 280 AA. AC Q25846_			
Dt	01-NOV-1996	(TREMBLrel. 01. Created)	
Dt	01-NOV-1996	(TREMBLrel. 01. Last sequence update)	
Dt	01-OCT-2002	(TREMBLrel. 22. Last annotation update)	
Dd	Liver stage-specific antigen 1 (Fragment).		
CN	Name=LSA-1;		
OS	Plasmodium falciparum.		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCB1_TaxID=583;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=96065165; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;		
RA	Yang C., Shi Y.P., Uchayakumar V., Alperts M.P., Povoa M.M.,		
RA	Hawley W.A., Collins W.E., Lal A.A.;		
RT	"Sequence variations in the non-repetitive regions of the liver stage-		
RT	specific antigen-1 (LSA-1) of Plasmodium falciparum from field		
RT	isolates";		
RL	Mol. Biochem. Parasitol. 71:291-294 (1995).		
DR	EMBL, L40887; AAC41597.1; -; Genomic DNA.		
FT	NON_TER	1	
SO	SEQUENCE	280 AA; 32882 MW; 46763641C81AFC33 CRC64;	
 Alignment Scores:			
Pred. No.:	1,21e-59	Length:	280
Score:	1442.00	Matches:	278
Percent Similarity:	99.64%	Conservative:	1
Best Local Similarity:	99.29%	Mismatches:	1
Query Match:	58.00%	Indels:	0
	2	Gaps:	0
 US-09-837-344-41 (1-1482) x Q25846_PLAFA (1-280)			
Oy	640	AGGAAGCGCTGATACGAAAAAAAATTTAGAAAAAGAAAAAGAAACATGAGATATATTAGCA	699
Dd	1	ArgLysAlaAerPtnLnrLysLysAsnLeuGluArgLysLysGluHisGlyAspAlaLeuAla	20
Oy	700	GAGGATTTATATGTCGTTTGAAGAAATACCAGSTATATGAAGAACTTCATCAGAAAAAGAACT	759
Dd	21	GluAspLeuTyrGlyArgLeuGluLnrLeuProAlaIleGluLeuProSerGluAsnGluArg	40
Oy	760	GGAATTTATATACCATCAATCTCTTAACTCTGAGACAACAGAGGGAATAGTAGAGAT	819
Dd	41	GlyLysTyrLnrLeuProHisGlnSerSerLeuProGlnAerPasnArgGlyAsnSerArgAsp	60

Oy		820	TCGAAGAAATCTTAATAATAGAAAAACAATAGAAGATCTATTACAACAAAATGTTGAA	879
Dd		61	SerLysGluIleSerLeilelleGIuYsrThraaaTgGluSerilleThrthrValGlu	80
Oy		880	GGACGAAAGGATATCATTAAGACACTCTTGAAAGAAAAGAAAGATCGTTCATTAACCA	939
Dd		81	GlyAArgAspRileHisLyseGlnHsbEugLuGLuLYsAspGlyserilleYsPro	100
Oy		940	GAACAAAAGAGATMAATCTGCTGACATACAAAATCATACATTAAGACAGCTAAATATT	999
Dd		101	GIuGIuLYsgLuAspLYserHlaAspRilleGIuaenHieThrLeuGIuThrValAsnIle	120
Oy		1000	TCTGATGTTAATGATTTTTCAAATAGTAAGTATGAGAGATGAATAAGTCTGAATATGC	1055
Dd		121	SerAspValaAsnaAspPheGlnIleSerLYsTYrGluAspGluIleSeralagLuYrAsp	140
Oy		1060	CATTCAATTAATAGTCAAGAAGATGATGAAGACTTAGACGAATTTAAAGCCTATTGCG	1111
Dd		141	AAlaserLeuIleaSpIuGIuAspAspGluAspAspGluAspGluAspGluAspGluAspGlu	160
Oy		1120	CAATATGACAAATTTCCAAAGATGAGAAAACATAGCAATTTAATAAGAACTAGAAGATTG	1177
Dd		161	GIuTYrAspAsnPhelnsPrgIuAspGluIleGIuIleTYrLYsGluLeuGIuAspGlu	180
Oy		1180	ATAGAGAAAAATGAAAAATTAGATGATTTAGATGAAGAAATGAATAATCATCAGAACAA	1233
Dd		181	IleGIuLYsaSGluAsnIleuAsnIleuAspIleuAspGluGIuIleGIuLYsSerSeGIuGIu	200
Oy		1240	TNATCTGAAGAAAAATTAATAAGAAAGAAATTAAGAAAAACAAGATATAATTTT	1299
Dd		201	LeuSeGIuGIuLYsIleLYsLYsGIuLYsTYrGluLYsThrLYsAspAsnPhne	220
Oy		1300	AAACCAATATGATAAAAAGTTTGATGATGAGCAATATTAATAATAATGAATGAAGCAG	1351
Dd		221	LYsProAsnAspLYsSerLeuTYrAspGluHsieIleLYsLYsTYrLYsAsnAspLYsGln	240
Oy		1360	GTTAATAGAGAAAAAGAAAAATTCATTAATATCTGTTTCATATATTTGACGAGACAAT	1411
Dd		241	ValaenLYsGluLYsGIuLYsPheIleLYsSerLeuPheHieIlePheAspGlyAspAsn	260
Oy		1420	GAATTTTACAGATCTCGATGATGATTAATCTGAAGATATTAATTAATTTATGAACATA	1477
Dd		261	GluIleLeuGlnIleValaAspGluLeuSeGIuAspRileThrLYsTYrPheMetLYsLeu	280
RESULT 13				
ID	Q25853_PLAFA PRELIMINARY;	PRT;	280 AA.	
AC	Q25853;			
DT	01-NOV-1996 (TREMBLrel). 01, Created)			
DT	01-NOV-1996 (TREMBLrel). 01, Last sequence update)			
DT	01-OCT-2002 (TREMBLrel). 22, last annotation update)			
DE	Liver stage-specific antigen 1 (Fragment).			
GN	Name=LSA-1,			
OS	Plasmodium falciparum.			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=5833;			
RN	(1)			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;			
RA	Yang C., Shi Y.P., Udhayakumar V., Alberts M.P., Povoa M.W.,			
RT	Hawley W.A., Collins W.E., Lai A.A.;			
RT	"Sequence variations in the non-repetitive regions of the liver stage-			
RT	specific antigen-1 (LSA-1) of Plasmodium falciparum from field			
RT	isolates."			
RL	Mol. Biochem. Parasitol. 71:291-294 (1995).			
DR	EMBL; L40835; AAB59234.1; -; Genomic_DNA.			
FT	NON TER			
SO	SEQUENCE	280 AA; 32938 MW; 46751C45F4DCBD3 CRC64;		
Alignment Scores:				
Pred. No.:	2, 08e-59	Length:	280	
Score:	1437.00	Matches:	277	
Percent Similarity:	99.29%	Conservative:	1	

US-09-837-344-41 (1-1482) x Q25853_PLAFA (1-280)	Best Local Similarity: 98.93%	Query Match: 57.80%	Matches: 2	Indels: 0	Gaps: 0
QY 640 AGGAAAGCTGCTACGAAAAAAATTTAGAAAAGAAAAAGAAACATTCGAGATATTTAGCA	699				
DB 1 ArglysaIaapRTrLyLyAsnLeuGluatrgLyAsnGluLysGlyAspValLeuLa	20				
QY 700 GAGGATTTATATGCGCTTTGAAATATCCAGCTTTAGAACTTCATGACGAAATATGAACTG	759				
DB 21 GluspleuTyrgLyAaGleuGluileProValilleGluLeuProSerGluAsnGluArg	40				
QY 760 GGATATTAATATATACATCAATCTTCTTTACCTGAGACAAACAGAGGAAATAGTAGAT	819				
DB 41 GlyTyrrTyrlleProHIsGlnSerSerLeuProGlnAspAaDgGlyAsnSerArgAsp	60				
QY 820 TCCAAAGAAATATCTATATATGAAAAAACAAATAGAAATCTTATTAACAACAAAGTTGAA	879				
DB 61 SerlysgluileSerilleIleGluLySthAsnArguSerilleTrHAsnValGlu	80				
QY 880 GGAGGAAAGGATATATACATTAAGACATCTTGAACAAAAAGAAAGATGGTCAATAAACCA	939				
DB 81 GlyArgAaAspilleHIsLySgLyHIsleuGluGluLySbAspGlySerilleLySPro	100				
QY 940 GAACAAAAAGAGATAAATCTGCTGACATACAAATCATACATTAGACAGTAATATT	999				
DB 101 GluGlnLySgLuAspLySereAlaAspilleGlnAsnHIsThrLeuGluThrValAsnille	120				
QY 1000 TCTGATGTTAATGATTTTCAAATAAGTAAGTATAGAGATGAAATAAGTCTGATATGAC	1055				
DB 121 SerAspValaAspAspHegInIleSerLyStryGluAspGluileSerAlaGluTyAsp	140				
QY 1060 GATTCATTAATATAGATGAAAGAAAGATGATGAAGCTTAGACGAAATTTAAAGCTATTTGTG	1119				
DB 141 AspSerLeuIleAspGluGluGlnAspAspGluAspLeuAspGluPheLySProIleVal	160				
QY 1120 CAATATGACAAATTTCCACAGATGAAGAAAAACATAGAAATTTATAAAGACTAGAAATTTG	1179				
DB 161 GlnTyrrAspAsnHegInAspGluGluAsnilleGlyIleTyrrLySgLuLeuGluAspLeu	180				
QY 1180 ATAGAGAAAAATGAAATTTTGAATGATTTAATGAAAGATATAGAAAAATCATCAGAGAA	1239				
DB 181 IleGluLySAsnGluAsnLeuAspAspLeuAspGluGlyIleGluLySereSerGluGlu	200				
QY 1240 TTATCTGAAAGAAAAATTAATAAAAAAGAAAGAAATATGAAAAACAAAGCATATTAATTT	1299				
DB 201 LeuSerGluGluLySleLySbLySgLyLySbLyStryGluLyStryrLySAspAsnHIsPhe	220				
QY 1300 AAACCAAAATGATTAAGATTTGATGATGACGATATTAATAAAAAATTAATAAATGATTAACG	1355				
DB 221 LySProAsnAspLySereLeuTyrrAspIleHIsIleLySbLyStryrLySAsnValLySgIn	240				
QY 1360 GTTAATTAAGAAAAAGAAAAATTCATTAATAATCATTTGTTCTATATTTGAAGAGACAAAT	1419				
DB 241 ValAsnLySgLyLySgLySbHelleLySereLeuPheHIsIlePheAspGlyAspAsn	260				
QY 1420 GAAATTTTACGATCGTCGATGAGATTAATCTGAAGATTAATACAAATATTTATGAAACTA	1479				
DB 261 GluIleLeuGlnIleValAspGluLeuSerGluAspIleHIsLyStryrPheMetLySLeu	280				
RESULT 14					
Q25854_PLAFA PRELIMINARY; PRT: 280 AA.					
AC Q25854; 01-NOV-1996 (TREMBLrel. 01, Created)					
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)					
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)					
DE Liver stage-specific antigen 1 (Fragment).					
GN Name=LSA-1					
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.					

NCBI_TaxID=5833;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Uchayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL/ L40836; AAC41604.1; -; Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32566 MW; 61851C5AD6E1A211 CRC64;

Alignment Scores:
Pred. No.: 2,86e-59 Length: 280
Score: 1434.00 Matches: 276
Percent Similarity: 99.29% Conservative: 2
Best Local Similarity: 98.57% Mismatches: 2
Query Match: 57.68% Indels: 0
DB: Gaps: 0

US-09-837-344-41 (1-1482) x Q25854_PLAFA (1-280)
OY 640 AGGAAGGCTGATACGAAAAAATTTAGAAAGAAAAAGAACATGAGATATATTGCA 699
DB 1 ArglysAlaAspThrLysLysAsnLeuGluArgLysGluHisGlyAspValLeuAla 20
OY 700 GAGGATTTATATGGCTTTAGAAATACAGCTATAGAACTTCATCAGAAATGAACGT 759
DB 21 GluAspLeuTyrGlyArgLeuGluIleProValIleGluLeuProSerGluAsnGluArg 40
OY 760 GGATATTATATATACCATCAATCTTCTTACCTCAGACAACAGAGGAATAGTAGAT 819
DB 41 GlyTyrTrileProHisGlnSerSerLeuProGlnAspAsnArgGlyAsnSerArgAsp 60
OY 820 TCCAGAGAAATATCTAATAAGAAAAACAATAGAGAACTTATTACAACTTAATTTGAA 879
DB 61 SerLysGluIleSerIleIleGluLysThrAsnArgGluSerIleThrAsnValGlu 80
OY 880 GGACGAAGGATATATCATTAAGAGACATCTTGAAGAAAGAAAGATGGTTCAATAAACA 939
DB 81 GlyArgArgAspIleHisLysGlyHisLeuGluGluLysLysAspGlySerIleLysPro 100
OY 940 GAACAAAAAGAGATAAATCTGCTGACATACAAAATCATACATTAGAGACATTAATTT 999
DB 101 GluGluLysGluAspLysSerAlaAspIleGlnAsnHisThrLeuGluThrValAsnIle 120
OY 1000 TCTGATGTTAATGATTTTCAAAATAGTAGAGATGAATAATTAAGCTGCAATATGAC 1059
DB 121 SerAspValAsnAspPheGlnIleSerLysTyrGluAspGluIleSerAlaGluTyrAsp 140
OY 1060 GATTCATTAATAGAGAGAAAGAGAGATGAGAGACTTAGAGCAATTTAGCTTATTTG 1119
DB 141 AspSerLeuIleAspGluGluGlnAspArgGluAspArgGluAspArgGluAspArgGlu 160
OY 1120 CAATATGACAAATTTCCAAAGTGAAGAAACAATAGCAATTTTAAAGAACTAGAGATTG 1179
DB 161 GlnTyrAspAsnPheGlnAspGluIleAsnIleGlyTyrLysGluLeuGluAspLeu 180
OY 1180 ATAGAGAAAAATGAAATTTAGATGATTTAGTAGAGAAATGAAGAAATCATCAGAGAA 1239
DB 181 IleGluLysAsnGluAsnLeuAspArgLysGluIleGluLysSerSerGluGlu 200
OY 1240 TTATCTGAAGAAAAAATAAAAAGAGAAAGAAATGAAAAACAAGATTAATTAATTT 1299
DB 201 LeuSerGluGluArgIleLysLysGlyLysTyrGluLysThrLysAspAsnAsnHe 220
OY 1300 AAACCAAAATGATAAAGTTTGTATGATGACATATTAATAAATATTAATAATGATTAAG 1359
DB 221 LysProAsnAspLysSerLeuTyrAspGlnHisIleLysLysTyrLysAsnValLysGln 240
OY 1360 GTTAATTAAGAAAAAGAAAAATTCATAAAATCATTTGTTTCAATATTTGACGAGACAA 1419

DB 241 ValAsnLysGluLysGluLysPheIleLysSerLeuPheHisIlePheAspGlyAspAsn 260
OY 1420 GAAATTTTACAGATCCGTGATGATGATATCTGAAGATTAATTAATTTTATGAACTA 1479
DB 261 GluIleLeuGlnIleValaAspGluLeuSerGluAspIleThrLysTyrPheMetLysLeu 280

RESULT 15
Q25845_PLAFA PRELIMINARY; PRT; 280 AA.
ID Q25845;
AC Q25845;
DT 01-NOV-1996 (TREMURel. 01, Created)
DT 01-NOV-1996 (TREMURel. 01, Last sequence update)
DT 01-OCT-2002 (TREMURel. 22, Last annotation update)
DS Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Uchayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL/ L40886; AAB59232.1; -; Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32908 MW; FDEB29F34DCAT7E9 CRC64;

Alignment Scores:
Pred. No.: 3.19e-59 Length: 280
Score: 1433.00 Matches: 276
Percent Similarity: 98.93% Conservative: 1
Best Local Similarity: 98.57% Mismatches: 3
Query Match: 57.64% Indels: 0
DB: Gaps: 0

US-09-837-344-41 (1-1482) x Q25845_PLAFA (1-280)
OY 640 AGGAAGGCTGATACGAAAAAATTTAGAAAGAAAAAGAACATGAGATATATTGCA 699
DB 1 ArglysAlaAspThrLysLysAsnLeuGluArgLysGluHisGlyAspValLeuAla 20
OY 700 GAGGATTTATATGGCTTTAGAAATACAGCTATAGAACTTCATCAGAAATGAACGT 759
DB 21 GluAspLeuTyrGlyArgLeuGluIleProValIleGluLeuProSerGluAsnGluArg 40
OY 760 GGATATTATATATACCATCAATCTTCTTACCTCAGACAACAGAGGAATAGTAGAT 819
DB 41 GlyTyrTrileProHisGlnSerSerLeuProGlnAspAsnArgGlyAsnSerArgAsp 60
OY 820 TCCAGAGAAATATCTAATAAGAAAAACAATAGAGAACTTATTACAACTTAATTTGAA 879
DB 61 SerLysGluIleSerIleIleGluLysThrAsnArgGluSerIleThrAsnValGlu 80
OY 880 GGACGAAGGATATATCATTAAGAGACATCTTGAAGAAAGAAAGATGGTTCAATAAACA 939
DB 81 GlyArgArgAspIleHisLysGlyHisLeuGluGluLysLysAspGlySerIleLysPro 100
OY 940 GAACAAAAAGAGATAAATCTGCTGACATACAAAATCATACATTAGAGACATTAATTT 999
DB 101 GluGluLysGluAspLysSerAlaAspIleGlnAsnHisThrLeuGluThrValAsnIle 120
OY 1000 TCTGATGTTAATGATTTTCAAAATAGTAGAGATGAATAATTAAGCTGCAATATGAC 1059
DB 121 SerAspValAsnAspPheGlnIleSerLysTyrGluAspGluIleSerAlaGluTyrAsp 140
OY 1060 GATTCATTAATAGAGAGAAAGAGAGATGAGAGACTTAGAGCAATTTAGCTTATTTG 1119
DB 141 AspSerLeuIleAspGluGluGlnAspArgGluAspArgGluAspArgGluAspArgGlu 160

```
Qy 1120 CAATATGACAAATTTCCAGATGGAAGAAAATTAAGAACTTAAGAACTAGAGATTG 1179
    |||
Db 161 GlnTyrAspAsnPhcGlnAspGlnuAsnIleGlyIleTyrLysGlnuLeuGlnuAspLeu 180
Qy 1180 ATAGAGAAAATGAAAATTTAGATGATTAGATGAAGAAATGAAAATCATCAGAGAA 1239
    |||
Db 181 ILeGlnLysAsnGlnuAsnLeuAspLeuAspGlnGlyIleGlnLysSerSerGlnu 200
Qy 1240 TTATCTGAAGAAAATTAAGAAAAGAAAGAAATATGAAAAACAAAGATATATATTT 1299
    |||
Db 201 LeuSerGlnGlnLysIleLysLysGlnLysLysTyrGlnLysThrLysAspAsnAspPhe 220
Qy 1300 AAACCAATGATTAAGTTTGTATGATGAGCATATTAATAATTAATAATGATAAGCAG 1359
    |||
Db 221 LysProAsnAspLysSerLeuTyrAspGlnuHisIleLysLysTyrLysAsnValLysGln 240
Qy 1360 GTTAATTAAGAAAAGAAAATTCATTAATAATCATTTGTTTCATATATTTTGACGAGACAAT 1419
    |||
Db 241 ValAsnLysGlnLysGlnLysPheIleLysSerLeuPheHisIlePheAspGlnAspAsn 260
Qy 1420 GAAATTTTACAGATCGTGATGAGTTATCTGAAGATATTAACATAATATTTTATGAAACTA 1479
    |||
Db 261 GluIleLeuGlnIleValAspGlnuLeuSerGlnuAspIleThrLysTyrPheMetLysLeu 280
```

Search completed: December 7, 2005, 06:55:28
Job time : 221.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 7, 2005, 03:41:06 ; Search time 140.5 Seconds
(without alignments)
9269.172 Million cell updates/sec

Title: US-09-837-344-41
Perfect score: 2486
Sequence: 1 CAGACACACACAGCATCT.....AATATTTATGAACTATA 1482

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p_model -DEV=xlh
-Q=/cgn2.1/USPRO.spool.h/US09837344/runat.05122005.133406.29453/app.query.fasta_1.1671
-DB=A-Geneseq -QFMT=fasta -SUFFIX=rag -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNIT5=bits -START=1 -END=1 -MATRIX=biom62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09837344 @CGN 1 1 183 @runat.05122005.133406.29453 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=130 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A-Geneseq.21.*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2478	99.7	493	2	AAR26944 P.falcip
2	1640.5	66.0	456	8	ADO21942 LSA-NRC(H
3	1630	65.6	457	8	ADO21920 LSA-NRC(H
4	1561	62.8	316	2	AAR26941 P.falcip
5	969.5	39.0	462	2	AAR05766 Porcion o
6	770.5	31.0	318	2	AAR26943 P.falcip
7	603	24.3	117	2	AAR26937 P.falcip
8	376.5	14.3	419	4	ABG17301 Novel hum
9	376	15.1	1898	2	AAY30795 A human c

10	376	15.1	1898	7	ADD48869	Ad448869 Human Pro
11	374	15.0	1439	8	ADR12189	Ad12189 Trichohya
12	374	15.0	1439	8	ADU77050	Adu77050 Murine ha
13	374	15.0	1439	8	ADM00404	Adm00404 Epithelia
14	373	15.0	1584	8	ABM80133	Abm80133 Tumour-as
15	361.5	14.5	412	2	AAW03626	Aaw03626 Human chy
16	359	14.4	1162	3	AAV96255	Aay96255 Kaposi's
17	359	14.4	1162	3	AAV96255	Aay96255 Kaposi's
18	359	14.4	1162	3	AAV96255	Aay96255 Kaposi's
19	359	14.4	1162	5	ABR05621	Abb05621 Kaposi's
20	359	14.4	1162	8	ADU55096	Adj55096 HIV8 late
21	359	14.4	1162	8	ADU55096	Adj55096 HIV8 late
22	339.5	13.7	207	3	AAV4672	Adv4672 Gene 33 h
23	335	13.5	562	2	AAV70491	Aav70491 Leucocyto
24	332.5	13.4	1192	3	AAV18165	Aab18165 Plasmodiu
25	331	13.3	2274	4	ABR58657	Abb58657 Drosophil
26	330.5	13.3	2990	8	ADP29884	Adp29884 Human sec
27	327.5	13.2	360	2	AAW03627	Aaw03627 Human fol
28	323.5	13.0	611	2	AAV29039	Aay29039 T. gondii
29	323.5	13.0	611	4	AAU25510	Aau25510 T. gondii
30	323.5	13.0	611	7	ADG17131	Adg17131 T. gondii
31	323	13.0	345	4	ABU53267	Abu53267 Human tes
32	320	12.9	658	4	ABR5632	Abb5632 Drosophil
33	320	12.9	800	4	ABR71459	Abb71459 Drosophil
34	320	12.9	3616	8	ADF45530	Adf45530 Chicken A
35	319	12.8	1972	2	AAW00024	Aaw00024 Smooth mu
36	318	12.8	554	4	ABR59454	Abb59454 Drosophil
37	317.5	12.8	1937	8	ADQ17289	Adq17289 Human bof
38	317.5	12.8	1937	8	ADQ17242	Adq17242 Human bof
39	316	12.7	1180	7	ADF06151	Adf06151 Bacterial
40	315.5	12.7	1170	5	ABB05710	Abb05710 Human tes
41	315.5	12.7	1170	7	ADC06789	Adc06789 Human pro
42	315.5	12.7	1170	7	ADC06845	Adc06845 Prostata
43	315.5	12.7	1501	6	ABP67998	Abp67998 Human pol
44	315.5	12.7	1564	6	ABP98812	Abp98812 Human str
45	315.5	12.7	1564	7	ADC06788	Adc06788 Human pro

ALIGNMENTS

RESULT 1	AAAR26944	standard; protein: 493 AA.
ID	AAAR26944	
XX	AAAR26944;	
AC		
XX		
DT	25-MAR-2003 (revised)	
DT	08-FEB-1993 (first entry)	
XX		
DB	P.falciparum LSA gene C-terminal region.	
XX		
KW	Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;	
KW	patulism; liver stage-specific antigen.	
XX		
OS	Plasmodium falciparum.	
XX		
FT	Key	Location/Qualifiers
FT	Region	13..213
FT		/label= repetitive_region
FT	Region	214..493
FT		/label= non-repetitive_region
XX		
PN	W09213884-A1.	
XX		
PD	20-AUG-1992.	
XX		
PR	05-FEB-1992; 92MO-FR000104.	
XX		
PR	05-FEB-1991; 91FR-00001286.	
XX		
PA	(INSP) INST PASTEUR.	
XX		
PI	Guerinmarchand C, Drulhe P;	

XX MPI: 1992-299985/36.
 DR N-PSDB; AAQ28119.
 XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
 PT vaccination against, treatment of and diagnosis of malaria.
 XX
 XX Claim 2: Fig 8-10; 81pp; French.
 XX
 CC The 3' part of the P. falciparum liver-stage specific antigen (LSA) gene
 CC codes for a polypeptide sequence which carries a T cell epitope
 CC characteristic of a protein produced in hepatocytes infected with
 CC P. falciparum. The polypeptide can be used in the preparation of vaccines
 CC against malaria. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 493 AA;

Alignment Scores:

Pred. No.:	2, 81e-177	Length:	493
Score:	2478.00	Matches:	493
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.68%	Indels:	0
DB:	2	Gaps:	0

US-09-837-344-41 (1-1482) x AAR26944 (1-493)

QY	1	CAGAAACAACAAGCGATCTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAAGACAA	60
DB	1	GIINGIINGIINseraspLeuGIINGIINargAlalysGIuIyBleuGIINGIIN	20
QY	61	CAAGCGATTTGAAACAAGATGACTTGCTAAAGAAAAGTTGCAAGAGCAAGCGAT	120
DB	21	GIINseraspLeuGIINsPargLeuAlalysGIuIyBleuGIINGIINserasp	40
QY	121	TTAGAACAGAGACGCTTCTAAAGAAAAGTTGCAAGACAAACAAGCGATCTAGAACAA	180
DB	41	IeugIINGIINargLeuAlalysGIuIyBleuGIINGIINseraspLeuGIIN	60
QY	181	GAGAGCGTCTTAAAGAAAAGTTGCAAGACAAACAAGCGATTTAGACAGAGACGT	240
DB	61	GIuIargAlalysGIuIyBleuGIINGIINseraspLeuGIINGIINargArg	80
QY	241	GCTTAAAGAAAAGTTGCAAGACAAACAAGCGATTTGCAAGACAAACAAGCGATTTG	300
DB	81	AlalysGIuIyBleuGIINGIINseraspLeuGIINsPargLeuAlalysGIu	100
QY	301	AACTTCAAGAGCAGCAAGCGATTTAGAACAGAGACGCTGCTAAAGAAAAGTTGCA	360
DB	101	IyBleuGIINGIINseraspLeuGIINGIINargArgAlalysGIuIyBleuGIIN	120
QY	361	GAAACAACAAGCGATTTAGAACAGAGACGCTGCTTAAAGAAAAGTTGCAAGACAA	420
DB	121	GIuIINGIINseraspLeuGIINGIINargAlalysGIuIyBleuGIINGIIN	140
QY	421	AGCGATTTAGAACAGAGACGCTTAAAGAAAAGTTGCAAGACAAACAAGCGATTTA	480
DB	141	SeraspLeuGIINGIINargLeuAlalysGIuIyBleuGIINGIINseraspLeu	160
QY	481	GAAACAAGAGCAGCTGCTTAAAGAAAAGTTGCAAGACAAACAAGCGATTTAGAACAG	540
DB	161	GIuIINGIINargArgAlalysGIuIyBleuGIINGIINseraspLeuGIINGIIN	180
QY	541	AGACGTGCTTAAAGAAAAGTTGCAAGACAAACAAGCGATTTAGAACAGAGACGTGCT	600
DB	181	ArgArgAlalysGIuIyBleuGIINGIINseraspLeuGIINGIINargArgAla	200
QY	601	AAAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAAAGCTGATAGAAAAA	660
DB	201	IyBGIuIyBleuGIINGIINargaspLeuGIINargIyBAlaapIrrIyBlys	220
QY	661	AAATTGAAAAGAAAAGAAATGAGATATATTAGACAGGATTTATGTCGTTTA	720

DB	221	AenIeugIINargIyBleuGIINargIyBleuAlalysGIuIyBleuGIINargLeu	240
QY	721	GAAATACCGCTATAGAACTTCCATCAGAAAATGAAAGCGATATATTATACCATGAA	780
DB	241	GIuIleProAlalysGIuIyBleuProserGIuIyBleuGIuIyBleuGIuIyBleu	260
QY	781	TCTTCTTACCTCAGACACAGAGGAAATAGTAGATTTCCAGAAAATCTTAAATA	840
DB	261	SeraspLeuProGIINsPargLeuAlalysGIuIyBleuGIINargIyBleuGIIN	280
QY	841	GAAAAACAATTAGAGATCTTATTCACAAATGTTGAAAGACGAAAGGATTTACATAA	900
DB	281	GIuIyBleuAlalysGIuIyBleuGIINargIyBleuGIINargIyBleuGIIN	300
QY	901	GGACATCTTGAAGAAAAGATGCTCAATTAACCAAGAAACAAGATTAATCT	960
DB	301	GIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleu	320
QY	961	GCTGACATACAAATCATATACATTTAGAGACAGTAAATATTTCTGATTTAATGATTTCA	1020
DB	321	AlaapIleGIINsPargLeuGIINargLeuGIINargLeuGIINargLeuGIIN	340
QY	1021	ATACTTAAGTATGAGATGAAATTAAGCTGCTGATTTGACGATTTCTTATAGATGAA	1080
DB	341	IleSerIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleu	360
QY	1081	GAAAGATGATGAAAGCTTATGAGCAATTTAAGCGCTATTTGCAATATGACATTTCCAA	1140
DB	361	GIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleu	380
QY	1141	GAAAGAAAATAGAAATTTATTAAGAACTAGAGATTTGATAGAGAAAATGAAATTTA	1200
DB	381	GIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleu	400
QY	1201	GATGATTTAGATGAAAGAAATCATATGAGAAATTTCTGAAAGAAAATTAATAA	1260
DB	401	AspaspLeuAspGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleu	420
QY	1261	AAAGGAAAATATGAAAGAAAATGAAAGTAAATTTTAAACCAATGATTAAGTTTG	1320
DB	421	IyBGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleu	440
QY	1321	TATGATGACATATTTAAATAATATTAATAATGATTAAGCAAGTTAATAGCAAGAAA	1380
DB	441	TyraspGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleu	460
QY	1381	TTTCATTAATCATTTGCTATATTTGACGAGACATGAAATTTTACAGATCGTGAT	1440
DB	461	PheIleIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleu	480
QY	1441	GAGTTATCTGAGATATTAATTAATTTTATGAACAATA	1479
DB	481	GIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleuGIuIyBleu	493

RESULT 2
 ADO21942
 ID ADO21942 standard; protein; 456 AA.
 AC ADO21942;
 DT 12-AUG-2004 (first entry)
 XX
 XX LSA-NRC(H) construct protein derived from Malaria parasite LSA-1.
 XX
 XX LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
 KW parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
 KW malaria parasite P. falciparum; LSA-NRC(H) construct.
 XX
 OS Plasmodium falciparum.
 OS Synthetic.
 XX
 XX MO200404167-A2.

KW parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
KM malaria parasite P. falciparum; LSA-NRC(H)Mut construct; mutant; mteain.
XX
OS Plasmodium falciparum.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 433 /note="Residue inserted relative to wild-type P.
FT falciparum T5 epitope"
XX
XX WO200404167-A2.
XX
XX 27-MAY-2004.
XX
XX 12-NOV-2003; 2003WO-US036011.
XX
XX 12-NOV-2002; 2002US-0425719P.
XX
XX (REED-) REED ARMY INST RES WALTER.
XX
XX Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W,
PI Barbosa A;
XX
XX WPI: 2004-420309/39.
XX
XX N-PSDB; ADO21919.
XX
XX
XX Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
XX
XX Claim 5; SEQ ID NO 4; 90pp; English.
XX
XX The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct protein of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to an amino acid insertion
CC within the T5 P. falciparum epitope and a C-terminal His6 tag.
XX
XX
SQ Sequence 457 AA;

Alignment Scores:
Pred. No.: 9, 566-114 Length: 457
Score: 1630.00 Matches: 338
Percent Similarity: 79.48% Conservative: 26
Best Local Similarity: 73.80% Mismatches: 70
Query Match: 65.57% Indels: 24
DB: 8 Gaps: 6

US-09-837-344-41 (1-1482) x ADO21920 (1-457)
QY 124 GAACAGAGAGAGACTTGAAGAAAAGTTGCAAGAACAAAGCATTTAGAACAGAG 183
DB 6 GTULYASPGIULIELYSESAENLEUASERGILYSESAEN-----SERA YG 23
QY 184 AGACGCTGAAGAAAAGTTGCAAGAACAAAGCATTTAGAACAGAGAGAGAGAGCTGCT 243
DB 24 ASNAIRGILEASNGILULYVSHIGILYLYSHIIVALLEUSERHISANSERTYRGIL 43
QY 244 AAAAGAAAAGTTGCAAGAACAAAGCATTTAGAACAGAGATGA----- 288
DB 44 LYSITHLEASHANGLUENANENLYSPHEPHEASPLYSGLULEUTHMETSER 63
QY 289 CTGTGAAGAAAAGTTGCAAGAGAGCAAGAGCATTTAGAACAGAGAGAGAGCTGCTAA 348
DB 64 ASNAVALYASNAVALSERGINTHIRANPHELYSERLEULEURGHANLEUGILVALSER 83

QY 349 GAAAAGTTGCAAGAACAAAGCATTTAGAACAGAGAGAGAGAGAGCTGCTAAAGAAAAGTTG 408
DB 84 GILUENILEPHEULYSGILUENLYSENLNENLYSGILUGILYSEULEUILEGLIHILE 103
QY 409 CAAGAACAAACAAAGCATTTGAAACAAAGAGACTTGTGAAGAAAAGTTGCCAGAACAA 468
DB 104 ILEASNPASPAEPAP-----LYSELYSYSTYRILEYSGILYGINASPGILUSNR YG 121
QY 469 CAAGCGATTTAGAACAAAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCAT 528
DB 122 GINGLUSPHEUGLU-----GILYSAIAIAGILUGINGLSERSEP 135
QY 529 TTAGAACAGAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGCATTTAGAACAA 588
DB 136 LEUGILUGINGILUARGLEULALYSGILYSEU-----GLN 147
QY 589 GAGAGAGCTGCTTAAGAAAAGTTGCAAGAGCAAGCAAGAAAGTTTGAACAAAGAGAGCT 648
DB 148 GILUARGLEULALYSGILYSEUGINGILUGINGILNARGAPLEUGILNARGYSAIA 167
QY 649 GATACGAAAAAATTTAGAAAAGAAAAGAACATGGAGATATATTTAGCAGAGATTTA 708
DB 168 ASPIHRLYSYASNLLEUGILUARGYSEYSGILYSEYVALLEULAGILUSPHEU 187
QY 709 TATGCTCGTTTGAAGAAATACAGCTATAGAACTTCATCAGAAATGAACGTGATATTAT 768
DB 188 TYRGILYAGLEUGILUILEPROLALILEGLILEUPROSERGILUASNGILUARGLYTYR YR 207
QY 769 ATACCAATCAATCTTCTTTACCTCAGAGCAACAGAGGAAATGTAAGAGATTCCAGAGAA 828
DB 208 ILEPROHISGINSERSEULEUPROGILNAPSHNRGILYASNSERVARXASPSERYSGIL 227
QY 829 ATATCTATATATAGAAAAAATAGAGATCTATTAACAACAAATGTTGAAGAGCAAGG 888
DB 228 ILESEITIEILEGLULYSTRASNRGILSEITIEHTRHASNVALGILUGILYRGARG 247
QY 889 GATATACATTAAGAGACATCTTGAAGAAAAGAAAGTGTTCATTAACCAAGAACAAA 948
DB 248 ASPILEHLSYBGLYHISLEUGILUGILYSEYSAAPGILYSEITILEYPROGILUGILN YS 267
QY 949 GAAGTAAATCTGCTGACATACAAAATCATCATTTAGAGACAGTAAATTTCTGATGTT 1008
DB 268 GIUASPLYSSEALASPDILEGLNASHNLSHRLSEUGILUTHVALASHNILESEXPVAL 287
QY 1009 AATGATTTCAAAATAGATAGATAGATGAGATGAATTAAGCTGAATATGACATTCATTA 1068
DB 288 ASNAEPHEGILILESELYSTYRGILUASPGILUILESEALALEGILTYRASHPASPERLEU 307
QY 1069 ATAGATGAAGAAAGATGATGAGACCTTAGACAAATTTAAGCTTATGCAATATGAC 1128
DB 308 ILEASPGILUGILUASPSAPGILUSPHEUASPGILUPHELYPROILEVALGINTRYAP 327
QY 1129 AATTTCCAAAGATGAGAAAACATACAGAAATTTAAGAACTTGAGAGATTGATAGAAA 1188
DB 328 ASNPHEGINASPGILUASNILEGLIYILEYTYRGLYSEUGILUASPHEUILEGLIUL YS 347
QY 1189 AATGAAAATTTAGATGATTTAGATGAGAAATTAAGAAATCATCAGAGAAATTTCTGAA 1248
DB 348 ASNGILUENLEUASPSAPGILUGILYILEGLULYSSERSEGLUGILUENUSERGIL 367
QY 1249 GAAAAAATTAAGAAAAGAAAAGAAATATGAAAAACAAAGATATATTTTAAACCAAT 1308
DB 366 GILYBILIELYBGLYLYSELYSTYRGILYSTRILYASPSANSHNPHELYSPROASH 387
QY 1309 GATAAAAGTTGTATGATGAGCATTTTAAAAATTTAAAAATGATPAGAGGTTAATAG 1368
DB 388 ASPLYSSELYTYASPGILUIHISILEYLSYSTYLYASASPSGLINVALASHNLYS 407
QY 1369 GAAAAGAAAATTCATTAATCATGTTTTCATATATTTGACGGAGACAATGAAATTTTA 1428
DB 408 GILYSGILYSPHEILEYSELEUPHEKSTILEPESAPGILYASPSANGLUILEU 427
QY 1429 CAGATCGTGGATGAG-----TTATCTGAAGATATATTAATATTTTAAAGAACTA 1479

Db 428 GlnIleValAspGluArgLeuSerGlnuSpIleThrIleTyPheMetuLysLeu 445

RESULT 4
AAR26941 ID AAR26941 standard; protein; 316 AA.

XX AAR26941;
AC 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX

DE P.falciparum LSA-R-NR protein.
KM Malaria; hepatocyte; sporozoite; plasmod DG 536; T-cell epitope;
KW paludism; liver stage-specific antigen.
XX

OS Plasmodium falciparum.
XX

FH Key Location/Qualifiers
FT Region 1..209
FT /label= repeat region
FT /note= "contains 12 x 17mer repeats"
FT 210..316
FT /label= non-repeat_region

PN M09213884-A1.
XX
XX 20-AUG-1992.
XX
XX 05-FEB-1992; 92MO-FR000104.
XX
XX 05-FEB-1991; 91FR-00001286.
XX
XX (INSP) INST PASTEUR.
XX
XX Guerimarchand C, Druilh P;
XX
XX WPI; 1992-299985/36.
DR N-PSDB; AA028115.
XX
XX
XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.
XX
XX
XX Disclosure; Fig 1; 81pp; French.

CC A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
CC gIII was used to transform E.coli. The expression library was screened
CC with human antisera against antigens of all stages of P. falciparum. The
CC library was rescreened with antibodies affinity- purified on a clone
CC which was able to recognise antibodies specific to the hepatic phase.
CC About 40 clones were detected which produced a characteristic LSA
CC epitope. The clone with the largest insert (950 bases) encoded LSA-R-NR
CC containing a 12-repeat region followed by a non-repeat region. Preferred
CC antigenic polypeptides of the invention are derived from the amino acid
CC sequence of LSA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX

Sequence 316 AA;
SQ

Alignment Scores:
Pred. No.: 1,356-108 Length: 316
Score: 1561.00 Matches: 316
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 62.79% Indels: 0
Gaps: 0

US-09-837-344-41 (1-1482) x AAR26941 (1-316)

QY 13 AGCGATCTGACACAGAGACGTCCTAAAGAAAGTTGCAAGACAAACAAACGATTTA 72
DB 1 SerAspLeuGluGlnGluArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeu 20

QY 73 GAACAAGATAGACTTGCTAAAGAAAGTTACAGACAGCAAGCGATTTAGAACAGAG 132
DB 21 GluGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGln 40

QY 133 AGACTTGCTAAAGAAAGTTGCAAGACAAACAAAGCGATCTTACAAACAGAGAGCTGCT 192
DB 41 ArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgAla 60

QY 193 AAAGAAAGCTTGCAAGACAAACAAAGCGATTTAGAACAGAGAGAGCTGCTAAAGAAAG 252
DB 61 LysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgAlaLysGluLys 80

QY 253 TTGCAAGACAAACAAAGCGATTTAGAACAGATAGACTTGCTAAAGAAAGTTACAGAG 312
DB 81 LeuGlnGlnGlnGlnSerAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGln 100

QY 313 CAGCAAGCGATTTAGAACAGAGAGAGCTGCTAAAGAAAGTTGCAAGACAAACAAAGC 372
DB 101 GlnGlnSerAspLeuGlnGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGlnGlnSer 120

QY 373 GATTTAGAACAGAGAGAGCTGCTAAAGAAAGTTGCAAGACAAACAAAGCGATTTAGAA 432
DB 121 AspLeuGlnGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGln 140

QY 433 CAAGAGAGCTTGCTAAAGAAAGTTGCAAGACAAACAAAGCGATTTAGAACAGAGAGAG 492
DB 141 GlnGlnArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGln 160

QY 493 CGTGCTAAAGAAAGTTGCAAGACAAACAAAGCGATTTAGAACAGAGAGAGCTGCTAA 552
DB 161 ArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgAlaLys 180

QY 553 GAAAGTTGCAAGACAAACAAAGCGATTTGCAAGAGAGAGAGCTGCTAAAGAAAGTTG 612
DB 181 GluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgAlaLysGluLysLeu 200

QY 613 CAAGAGAGCAAGAGATTTAGAACAAAGAGAGCTGATACGAAAAAATTTAGAAAGA 672
DB 201 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220

QY 673 AAAAGGACATGAGATATATTAGACAGAGATTTATGTCGTTTGAATTCACGCT 732
DB 221 LysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240

QY 733 ATGAACTTCATCAGAAATGAACCTGATATTATTATACCAATCAATCTTTCATCCT 792
DB 241 IleGlnLeuProSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260

QY 793 CAGAGACACAGAGGATTTAGAGATTTCCAGAAATATCTATATAGAAAAACAAT 852
DB 261 GlnAspAsnArgGlnAsnSerArgAspSerLysGlnIleSerIleIleGlnLysThrAsn 280

QY 853 AGAAGATCTATTACAAATGTTGAGAGAGAGGATTTATACATAAGACATCTTGAA 912
DB 281 ArgGlnSerIleThrThrAsnValGlnGlnArgArgAspIleHisLysGlnHisLeuGln 300

QY 913 GAAAAAGAAAGATGTTCAATPAAACCAAGACAAACAAACAAAGATTAATCT 960
DB 301 GluLysLysAspGlnSerIleLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 316

RESULT 5
AAR05766 ID AAR05766 standard; protein; 462 AA.

XX AAR05766;
AC 25-MAR-2003 (revised)
DT 05-NOV-1990 (first entry)
XX

DE Portion of peptide antigen to malarial sporozoite.
KW Malaria; sporozoite; vaccine; exoerythrocytic parasites; tetanus toxoid.
XX

[illegible]

Db	234	GIUcInuSprrgLeuAlaLySGluLyLeuGIngluGInGInSerAspLeuGInuSp	253
Oy	541	AGAGCTGCTAAAGAAAAGTTGCAGAAACAACAAAGCATTATTAAGAAACAAGAGACGCGCT	600
Db	254	ArgArgAlaLySGluLyLeuGIngluGInGInSerAspLeuGInuSp	273
Oy	601	AAAAAAAAGTTGCAGAGCAGCAAGAGATTTAAGACAAAGAGAGCGCTGATACCAAAAAA	660
Db	274	LySGluLyLeuGInGInGInGInSerAspLeuGInuGrrLySAlaSerThrGluThr	293
Oy	661	AATTAGAAAGAAAAAGAAACATGAGATATATTACAGAGAGATTATATGTCGTTTA	720
Db	294	LeuGIngluArg	297
Oy	721	GAAATTACCAGCTATGAACTTCCATCAGAAAAATGAAGTGGATTTATATACCAATCAA	780
Db	298	-----Gln	298
Oy	781	TCTTCTTACTCAGACCAACAGAGGGAAATGTAAGATTCCAGAAATATCTATATA	840
Db	299	SerAspLeuGIngluArgAlaLySGluLyLeuGIngluGInGInSerAspLeu	318
Oy	841	AAAAAAACAAATAGAGATCTATTACAAACAAATGTTGAAGAGCAGAGGATATCATAA	900
Db	319	GIUcIngluArgAlaLySGluLyLeuGIn-----GluGInGInuArgAspLeuGIn	337
Oy	901	GGACATCTTGAAGAAAAAGAAAGATGCTTCAATAAACCAAGAACAAAAAGAGTAACT	960
Db	338	GIuArgLeuAlaLyS-----GluLyLeuGInGInGIn	349
Oy	961	GCTGACATACAAATTCATACATTAGAGACAGTAATATTTCTGATGTTATGATTTCAA	1020
Db	350	SerAspLeuGIngluArgLeuAlaLySGluLyLeu-----	362
Oy	1021	ATAAGTAAGTATGAGATGAATAAGTCGTGAATATGACGATCATTAATAGATGAAGAA	1080
Db	363	-----Gln	363
Oy	1081	GAAGATGATGAAGACTTAGACGAATTTAAAGCCTATTGTCGAATATGACAAATTTCCAAGAT	1140
Db	364	GIUcIngluArgAspLeuGIngluArg-----LeuAlaLySGluLyLeuGIn	381
Oy	1141	GAAAGAAACATAGGAATTTATTAAGAACTAGAAAGTTGATAGAGAAAAATGAATAATTTA	1200
Db	382	GInGInSer-----AspLeuGIngluArgAlaLySGluLyLeu	396
Oy	1201	GATGATTTAGATGAAGATAGAAAAAATCATCAGAGATTTATCTGAAGAAAAATAAAA	1260
Db	397	GIUcIngluGInSerAspLeuGIn-----GluArgLeuAlaLySValLyLeuGIn	414
Oy	1261	AAAGAAACAAATAT--GAAAAACAAG-----GATTAATATTTTAAACCAAAATGAT	1311
Db	415	GIUcInGInSerTyrLeuGInuArgThrLySAlaSerThrGluThrLeuGIngluArgGln	434
Oy	1312	AAAGCTTGTTGATGAGCATATTTAAAAAATATAA---AATGATAGCAGGTTAATAG	1366
Db	435	SerAspLeuGIngluArgLeuAlaLySGluLyLeuGIngluGInGInSerAspLeu	454
Oy	1369	AAAAAGAAAAATTCATATAA	1389
Db	455	GIUcIngluArgArgAlaLyS	461
RESULT 6			
AAR26943			
ID AAR26943 standard; protein; 318 AA.			
XX AAR26943;			
XX			
DT 25-MAR-2003 (revised)			
DT 08-FEB-1993 (first entry)			
XX			
DE P.falciparum LSA N-terminal portion.			
XX			

KM Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
KW paludism; liver stage-specific antigen; ss.
OS Plasmodium falciparum.
XX
XX Key Location/Qualifiers
FH Region 1..153
FT Region /label= unique_5'_region
FT Region 154..318
FT Region /label= repetitive_region
XX
XX WO9213884-A1.
XX
XX PD 20-AUG-1992.
XX
XX PF 05-FEB-1992; 92WO-FR000104.
XX
XX PR 05-FEB-1991; 91FR-00001286.
XX
XX PA (INSP) INST PASTEUR.
XX
XX PI Guertimarchand C, Drullhe P;
XX
XX DR WPI; 1992-299985/36.
XX
XX DR N-PSDB; AAQ28117.
XX
XX PT Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.
XX
XX PS Claim 11; Fig 7; 81pp; French.
XX
XX CC The 5' part of the P.falciparum liver-stage specific antigen (LSA) gene
CC codes for a polypeptide sequence which carries a T cell epitope
CC characteristic of a protein produced in hepatocytes infected with
CC P.falciparum. The polypeptide can be used in the preparation of vaccines
CC against malaria. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 318 AA;
SQ
Alignment Scores:
Pred. NO.: 2.28e-49 Length: 318
Score: 770.50 Matches: 163
Percent Similarity: 92.59% Conservative: 12
Best Local Similarity: 86.24% Mismatches: 13
Query Match: 30.99% Indels: 1
DB: 2 Gaps: 1
US-09-837-344-41 (1-1482) x AAR26943 (1-318)
QY 16 GATCTAGAACAGAGACGCTGCTAAAGAAAAAGTTGCAGAAACAACAAGCGATTAGAA 75
DB 131 Aegpseapelyslylyetyrillelyeglylnaepgluaenaargglnlualapleuglu 150
QY 76 CAAGTACACTTGCTTAAAGAAAAAGTTACAAGACGCAAAACGATTAGAACAGAGAGA 135
DB 151 ---gluysalaaalysglulysleuglnglnglnserapsersegluglngluarg 169
QY 136 CTGCTTAAAGAAAAAGTTGCAGAAACAACAAGCGATTAGAACAGAGAGAGCGTTAAA 195
DB 170 Argalalyeglyulysleuglnglnglnserapsersegluglngluargleualalye 189
QY 196 GAAAAAGTTGCAGAAACAACAAGCGATTAGAACAGAGAGAGCGTTAAAGAAAAAGTTG 255
DB 190 Glulysleuglnglnglnglnserapsersegluglngluargalalyeglyuleu 209
QY 256 CAAGAACACAAGCGATTAGAACAGATAGCTTGTAAAGAAAAAGTTACAAGAGAG 315
DB 210 Glnlgluglnglnserapsersegluglngluargleualalyeglyuleuglnglu 229
QY 316 CAAGCGATTAGAACAGAGAGAGCGTCTAAAGAAAAAGTTGCAGAAACAACAAGCGAT 375
DB 230 Glnserapsersegluglnglnglnserapserseglulysleuglnglnglnglnserap 249

QY 376 TTAGAACAGAGAGACGCTGCTTAAAGAAAAAGTTGCAGAAACAACAAGCGATTAGAACAA 435
DB 250 Leuglnglnglnglnlunrgrharglalyeglyuleuglnglnglnglnserapleuglnglu 269
QY 436 GAGAGACTTGCTTAAAGAAAAAGTTGCAGAAACAACAAGCGATTAGAACAGAGAGAGCT 495
DB 270 Gluargleualalyeglyuleuglnglnglnglnserapsersegluglngluargleu 289
QY 496 GCTTAAAGAAAAAGTTGCAGAAACAACAAGCGATTAGAACAGAGAGAGCGTCTAAAGAA 555
DB 290 Alalyeglyuleuglnglnglnglnserapsersegluglnglnglnlunrgrharglalyeglu 309
QY 556 AAGTGCAGAGAACACAAGCGATTTA 582
DB 310 Argleuglnglnglnglnlunrgrhargleu 318
RESULT 7
AAR26937
ID AAR26937 standard; peptide; 117 AA.
XX
XX AC AAR26937;
XX
XX DT 25-MAR-2003 (revised)
XX
XX DT 08-FEB-1993 (first entry)
XX
XX DE P.falciparum liver-stage specific antigen non-repeat region.
XX
XX KW Malaria; LSA-NR; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;
XX paludism.
XX
XX OS Plasmodium falciparum.
XX
XX FN WO9213884-A1.
XX
XX PD 20-AUG-1992.
XX
XX PF 05-FEB-1992; 92WO-FR000104.
XX
XX PR 05-FEB-1991; 91FR-00001286.
XX
XX PA (INSP) INST PASTEUR.
XX
XX PI Guertimarchand C, Drullhe P;
XX
XX DR WPI; 1992-299985/36.
XX
XX PT Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.
XX
XX PS Claim 4; Page 55; 81pp; French.
XX
XX CC A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
CC gtl1 was used to transform E.coli. The expression library was screened
CC with human antisera against antigens of all stages of P.falciparum. The
CC library was rescreened with antibodies affinity-purified on a clone
CC which was able to recognise antibodies specific to the hepatic phase.
CC About 40 clones were detected which produced a characteristic LSA
CC epitope. The clone with the largest insert encoded LSA-R-NR containing a
CC 12-repeat region (LSA-R) followed by a non-repeat region (LSA-NR). The
CC sequence given here is a preferred polypeptide of the invention carrying
CC a T cell epitope typical of liver-stage P.falciparum. See AAR26919-
CC R26933, AAR26937 and AAR26939. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
XX SQ Sequence 117 AA;
SQ
Alignment Scores:
Pred. NO.: 7.16e-37 Length: 117
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.26% Indels: 0
DB: 2 Gaps: 0

US-09-837-344-41 (1-1482) x AAB26937 (1-117)

QY 610 TTGCAAGACGACCAAGAGATTAGAACCAAGAGCGTATACGAAAAAATTTAGAA 669

DB 1 LeuGInGluGInGlnAhrGAspLeuGInGlnAhrGAspThrLysLysAsnLeuGln 20

QY 670 AAAAAAGGACATGAGATATATTAGACAGAGATTTATGTCGTTTAAAGATACCA 729

DB 21 AArgLysLysGlnHISGlyAspLLeuValAGluAspLeuTyrGlyArgLeuGlnLeu 40

QY 730 GCTATGAACTTCATCAGAAATGAAACGTGATATTATATACCAATCAATCTCTTTA 789

DB 41 AAlIeGluLeuProSerGluAsnGlnAArgGlyTyrTyrLleProHISGlnSerLeu 60

QY 790 CCTCAGACACACAGAGGAAATGTAGAGATTTCCAGAAATTCATTAATAGAAAAACA 849

DB 61 ProGlnAspAsnAhrGlyAsnSerAhrGAspSerLysGlnLleSerLleGluLysThr 80

QY 850 AATAGAGATCTATTACACAAATGTGAAGACGAGAGGATATCATTAAGACATCTT 909

DB 81 AAsnAhrGluSerLleThrThrsnValGluGlyAhrGAspLleHISLysGlnHISLeu 100

QY 910 GAAGAAAAGAAAGATGTTCAATTAACCAAGAACAAAGAAAGATTAATCT 960

DB 101 GluGlnLysLysAspLysSerLleLysProGluGlnLysGluAspLysSer 117

RESULT 8

ABG17301

ID ABG17301 standard; protein; 419 AA.

XX ABG17301;

AC 18-FEB-2002 (first entry)

XX 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #17292.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

OS WO200175067-A2.

XX 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYTE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YF,

PI MPI: 2001-639362/73.

DR N-PSDB; AAS81488.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 47660; 103bp; English.

PS The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

XX Sequence 419 AA;

XX

Alignment Scores:

Pred. No.: 7,81e-20 Length: 419

Score: 376.50 Matches: 149

Percent Similarity: 44.34% Conservative: 35

Best Local Similarity: 35.90% Mismatches: 76

Query Match: 14.26% Indels: 155

DB: 4 Gaps: 25

US-09-837-344-41 (1-1482) x ABG17301 (1-419)

QY 1217 CCTCATCAATCAATCAATTAATTTCTCTATCAATCTTCTAGTCTTTATTA 1158

DB 136 ProAlaProAlaSerSerProCysSer--ArgSerLeuSerSerAhrSerLeu--- 153

QY 1157 ATTCTATGTTTCTTCATCTTGAA-----TTGTCATATTGCA 1116

DB 154 -----SerSerSerCysAhrGValSerCysCysHISerLeuSerSerCysCys 169

QY 1115 ATAGCTTAATTCCTTAAGTCTTCAATCATCTTCTTCTCATCTTAATTAATGCTCA 1056

DB 170 ArgSerLeuSerCysSerAhrSerSerAhrSerAhrSerAhrSerAhrSerSer 189

QY 1055 TATTCAGACATTATTCATCCCATCTTACTATTGTAATCAATTAACATCAGAAATA 996

DB 190 SerSerLeuSerLeuSerSerSerLeu----- 198

QY 995 TTTACTGCTCAATAGATGATTTGTTATGTCAGAGATTATCTTTTGTCTGCT 936

DB 199 -----SerCysSerSerCysAhrGLeuSer-----CysSer--- 208

QY 935 TTTATTAACCATCTTCTTCTTCTTCAAGATGCTTTATGTAATCCCTTGCTCTCA 876

DB 208 ----- 208

QY 875 ACATTGTTGTAATAGATTCTCTATTGTTTTTCTATTATAGATTATTCCTGGAATCT 816

DB 208 ----- 208

QY 815 CTACTATTCCCTGTGTGCTGAGTAAGAAGATTGATATATATATCCACGT 756

DB 209 -----Asn 209

QY 755 TCATTTTCTGATGAGATTCTATAGCTGATTTCTTAACGACCATATTAATCTCTGCT 696

DB 210 SerPheSerThrSerSerSerLeuSerLeuSer----- 221

QY 695 AATATATATCCCATGTTCTTTTCTTCTTAATTTTTTTTGTATACAGCTTCTCTTGT 636

DB 222 -----SerSerLeuSerCys 226

QY 635 TCTAATCTCTTGTGCTCTGCAACTTTCTTAGACAGCTCTCTGTTCTTAATG 576

DB 227 Ser---SerCysCysLeuSerCysSerAhrSerPheSerThrSerSerSerLeuSer 245

QY 575 CTTTGTTGTTCTTGAACATTCTTATGACAGCTCTCTCTGTTCTTAATGCTGTTGTT 516

DB 246 LeuSerSerSerLeuSerCysSerSerCysAhrGLeuSerCysSerAhrSerPhe----- 263

QY	673	AAAAAGACACATGAGATATATTATGACAGAGATTATATGATCGT-----	717
Db	794	LyseRgluaRgLy-----ArgInlRgLeuSerAlaRgProLeuRglu	810
QY	718	-----TTAGAAATACACGCTATAGAACTTCATCAGAAAAATGAACGTGATATTAT	768
Db	811	GlInRgLuARgGlInLeuARglAgluInuRgInglInARgGlInuRg-----Phe	828
QY	769	ATACCACATCAATCTTCTTTACCTCGACGACAAACAGAGGAATAGTAGAATCCAGGAA	828
Db	829	LeuProGluGluGluGluGluLeuSgluInARgLyARgGlInARgLyARgLuARgLuInuSglu	848
QY	829	ATATCTATTAATAGAAAAACAATAAG-----GAATCTATTGCA	867
Db	849	LeuGlInPheLeuGluGluGluGluGluInLeuGlInARgLyGluARgAlaGluGluInleuGlu	868
QY	868	ACAAATGTTGGAAGA---CGAAGGATATACATMAAGACATCTTGAAAGAAAAGAAAGAT	924
Db	869	GlUglUglUaRgPglYleuGluGluuARgIngluARgARgARgGluGluInARgARgR	888
QY	925	GATTCAATPAAAACCAACAAAAGAAAGATAATTCGCTGACATACAAATCATCATTA	984
Db	889	GlInuYstrPARgTyrGluInleuGluGluInuARgLyARg-----ARgARhIethreu	906
QY	985	GAGACAGTAATATTTCGATGCTTATGATGTTTCAAAATAGTAAGATAGAGAGTGAATA	1044
Db	907	-----TyrAlaLyRProAlaLeuGluGluInleu	916
QY	1045	AGTGTGAAATATGACGATTCATTATATAGATGAAGAAAGATGATGAAGACTTAGACGAA	1104
Db	917	ArgLySglu-----GluGluInleuGluGluGluGluGluGluGluInleuGluInARgGlUglU	934
QY	1105	TTTAAAGCTTATGTCCAATATGACATTTCCAAAGTGAAGAAACATAGGAATTTATATA	1164
Db	935	-----ArgGlUuYsARgARgARgGluGluGluGluInuARgGlInuYrARg	948
QY	1165	GAACTAGAAAGATTGATAGAGAAAATGAATAATTATAGATGATGAGAAATAGAA	1224
Db	949	GlUglUglUglUInleuGluGluGluInleuGluGluGluInleu-----LeuARgGlUglUARgGlU	966
QY	1225	AAATCATCAGAAATTAATCTGAGAAAAAATMAAAAAAGAAAGAAATATGAAAAACA	1284
Db	967	LyseARgARgARgGluInuARgGlInuARgGluInuYrARgLySARpLySleuGluGluInu	986
QY	1285	AAAGCAATATATTT---AAACCAATGATATAAGTTGTATGATGACATATTAAAAA	1341
Db	987	GlUglUglUInleuLeuGluGluGluInuProGluYsARgARgARgGluGluInuARgGlUuYs	1006
QY	1342	TATTAATAATGATAGACGTTATATATAGAAAAAGAAAAATTCATAAA	1389
Db	1007	TyrARgGlUglUglUglUInleuGluGluGluGluGluGluGluInleuLeuARg	1022
RESULT 10			
ID	ADD48869	standard, protein; 1898 AA.	
AC	ADD48869;		
DT	02-DEC-2004	(rev1ased)	
DT	29-JAN-2004	(first entry)	
DE	Human Protein AAA65582, SEQ ID NO 14580.		
XX	Human; pain; neuronal tissue; gene therapy;		
KW	spinal segmental nerve injury; chronic constriction injury; CCI;		
KW	spared nerve injury; SNi; Chung.		
OS	Homo sapiens.		
OS	Unidentified.		
PN	WO2003016475-A2.		

```

PD      27 -FEB-2003.
XX
XX      14-AUG-2002; 2002MO-US025765.
PF
PR      14-AUG-2001; 2001US-0312147P.
XX
PR      01-NOV-2001; 2001US-0346382P.
XX
PR      26-NOV-2001; 2001US-0333347P.
XX
PA      (GEHO ) GEN HOSPITAL CORP.
XX
PA      (FARB ) BAYER AG.
XX
PI      Woolf C, D'urso D, Befort K, Costigan M;
DR      MPI, 2003-268312/26.
XX
DR      GENDANK; AAA65582.
PT
PT      New composition comprising two or more isolated polypeptides, useful for
XX      preparing a medicament for treating pain in an animal.
XX
PS      Example 1; Page; 1017pp; English.
XX
XX      The invention discloses a composition comprising two or more isolated rat
XX      CC or human polynucleotides or a polynucleotide which represents a fragment,
XX      derivative or allelic variation of the nucleic acid sequence. Also
XX      CC claimed are a vector comprising the novel polynucleotide, a host cell
XX      CC comprising the vector, a method for identifying a nucleotide sequence
XX      CC which is differentially regulated in an animal subjected to pain and a
XX      CC kit to perform the method, an array, a method for identifying an agent
XX      CC that increases or decreases the expression of the polynucleotide sequence
XX      CC that is differentially expressed in neuronal tissue of a first animal
XX      CC subjected to pain, a method for identifying a compound which regulates
XX      CC the expression of a polynucleotide sequence which is differentially
XX      CC expressed in an animal subjected to pain, a method for identifying a
XX      CC compound that regulates the activity of one or more of the
XX      CC polynucleotides, a method for producing a pharmaceutical composition, a
XX      CC method for identifying a compound or small molecule that regulates the
XX      CC activity in an animal of one or more of the polypeptides given in the
XX      CC specification, a method for identifying a compound useful in treating
XX      CC pain and a pharmaceutical composition comprising the one or more
XX      CC polypeptides or their analogues. The polynucleotide or the compound that
XX      CC modulates its activity is useful for preparing a medicament for treating
XX      CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX      CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX      CC therapy). The sequence presented is a human protein (described in Table 3
XX      CC of the specification) which is differentially expressed during pain.
XX      CC Note: The sequence data for this patent did not form part of the printed
XX      CC specification, but was obtained in electronic form directly from WIFO at
XX      ftp.wifo.int/pub/published_pct_sequences.
SQ      Sequence 1898 AA;

Alignment Scores:
Pred. No.:          1,02e-19              Length:          1898
Score:              376.00                 Matches:           137
Percent Similarity: 47.33%                  Conservative:     117
Best Local Similarity: 25.56%                Mismatches:       174
Query Match:        15.12%                   Indels:           108
DB:                7                       Gaps:             24

US-09-837-344-41 (1-1482) x ADD48869 (1-1898)
QY      1 CAAGAACAAACAAGCGATCTTAGAACAACAGAGACGTCGTAAAGAAAAGTTGCACAGAACA 60
         ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      522 ArGGlUGlNgInleuArGrGlUGlNgInGUlUArGArGluGlnArgleuWybAtgBgIn 541
         ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      61 CAAAGCGATTAGAACCAAGATAGACTTGTCTAAAGAAAAAGTTACACAGACGACCAAAGCAT 120
         ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      542 GlU-----GlUGlUGlUArgleuGInGInArgleuArgserGIUGlNGInleuArg 558
         ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      121 TTAGAACAA---CAGAGACTTGCTATAAGAAAAGTTCCAGGAACAACAAGCGATCTTAGAA 177
         ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      559 ArGgIUGlNgInglUArgleuNlgInleuLeuLYbArGglUGlUGlUblUglWybAtgBleuGu 578

```

[illegible]

QY 1357 CAGGTTAATAGAAAGAA 1377
Db 1141 Lyspnearglgluglugin 1147
RESULT 13
ADW00404
ADW00404 standard; protein; 1439 AA.
XX
XX ADW00404;
XX
XX 07-APR-2005 (first entry)
XX
XX Epithelial hair follicle related antigen, SEQ ID 157.
XX
XX Endocrine-Gen.; Cardiovascular-Gen.; Vunerary; Osteopathic; Cardiant;
XX Gene Therapy; antibody production; antibody therapy; antigen;
XX hair growth; arteriosclerosis; osteoporosis; bone injury; angina; burns;
XX wound healing.
XX
XX Mus musculus.
XX
XX US2005008602-A1.
XX
XX 13-JAN-2005.
XX
XX 07-JAN-2004; 2004US-00754079.
XX
XX 05-JUN-2000; 2000JP-00166903.
XX 04-JUN-2001; 2001WO-JP004691.
XX 13-NOV-2001; 2001JP-0034738.
XX 13-NOV-2001; 2001JP-00347340.
XX 05-DEC-2001; 2001JP-00371175.
XX 05-DEC-2001; 2001JP-00371366.
XX 23-MAY-2002; 2002US-00155922.
XX 08-JAN-2003; 2003JP-00001891.
XX
XX (SUME) SUMITOMO ELECTRIC IND LTD.
XX
XX Hirai Y, Takebe K;
XX
XX WPI; 2005-080487/09.
XX
XX N-PSDB; ADW00406.
XX
XX New antibody specifically recognizing an antigen of 220kDa present in
XX epithelial new follicles, useful for evaluating and inducing hair growth
XX promoting activity and morphogenesis.
XX
XX Claim 5; SEQ ID NO 157; 87pp; English.
XX
XX The present invention relates to a novel antibody, or its fragment, which
XX specifically recognizes an antigen of about 220kDa (ADW00404) present in
XX new epithelial hair follicles. The antigen is specifically expressed
XX during the growth period of an imago or the developing period of a fetus.
XX Also claimed is a method for the evaluation of hair growth promoting
XX activity, comprising incubating skin tissue derived from a mammal in the
XX presence of a substance to be tested to promote hair growth, recovering
XX the skin tissue, reacting the skin tissue with the antibody, cited above,
XX and detecting the antibody, or its fragment, that reacts with the skin
XX tissue. The method and antibody of the invention are useful for
XX evaluating and inducing hair growth promoting activity, treating and
XX ameliorating symptoms of diseases associated with abnormal morphogenesis,
XX inducing cardiovascular revascularization, regeneration and endothelial
XX cell growth, such as in arteriosclerosis, Berger's disease, osteoporosis,
XX fractures and angina pectoris, and in treating burns or wounds.
XX
XX
XX Sequence 1439 AA;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1,4e-19 Length: 1439
XX Score: 374.00 Matches: 143
XX Percent Similarity: 48.15% Conservative: 130
XX Best Local Similarity: 25.22% Mismatches: 170
XX Query Match: 15.04% Indels: 124

DB: 9 Gaps: 25
US-09-837-344-41 (1-1482) x ADW00404 (1-1439)
QY 1 CAAGAACAAAGCGATCTTGAAACAAAGAGACGCTGCTAAGAAAG----- 48
Db 597 GluGlugluLysbArggluLeuGlugluLgluLuarArgbArgbGlnArgbArgbGln 616
QY 49 ---TTGCAAGAACAAAGCGATTTAGAACAAAGATTCGCTAAA-----GAA 96
Db 617 IleuGlugluLgluLgluLpheGlnArgGlnuLseGlnArgGlnuLArgbArgbGln 636
QY 97 AAGTTACAAAGACGCA----- 114
Db 637 ThrpheGlnGluGluGluLgluLleuGlnGlyLseRArgbArgbGlnGlnuArgbGln 656
QY 115 ---ACGATTTAGAACAAAGACCTTCTAAGAAAGTTGCAAGAACAAAGCGAT 171
Db 657 GlyLysPheLeuGlugluLgluLuarGlnLeuArgThrgLuarArgbGlnuArgbArg 676
QY 172 CTAGAACAAAGAGACGCTGCTAAGAAAG-----TTGCAAGAACAAAGCGAT 222
Db 677 GlnGluGlnGlnuLuarGlnPheGlnGlnGlnuLgluLgluLseGlnGlnuArgbGln 696
QY 223 TTAGAACAAAGAGACGCTGCTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAA 282
Db 697 LeuArgGlnGlnuLcybArgbArgbArgbArgbGlnGlnuArgbArgbGlnuArgbGln 716
QY 293 GATAGACTTCTTGAAGAAAGTTACAAAGCGCAAGCGATTTGAAACAAAGAGA--- 339
Db 717 GluGlnLeuArgbArgbGlnGlnu---ArgbArgbGlnuArgbArgbGlnGlnuArgbGln 735
QY 340 ---CGTCTAAGAAAGTTGCAAGAACAAAGCGATTTGAAACAAAGAGA--- 390
Db 736 LeuGlnuArgbGlnGlnuLgluLleuArgbArgbArgbArgbArgbGlnGlnuArgbGln 755
QY 391 ---CGTCTAAGAAAGTTGCAAGAACAAAGCGATTTGAAACAAAGAGA--- 441
Db 756 GlnGlnuArgbGlnGlnuLgluLleuArgbArgbArgbArgbArgbGlnGlnuArgbGln 775
QY 442 ---CTTGCTAAGAAAGTTGCAAGAACAAAGCGATTTGAAACAAAGAGA--- 492
Db 776 GlnGlnuArgbGlnGlnuLgluLleuArgbArgbArgbArgbArgbGlnGlnuArgbGln 795
QY 493 ---CGTCTAAGAAAGTTGCAAGAACAAAGCGATTTGAAACAAAGAGA--- 543
Db 796 GlnGlnuArgbGlnGlnuLgluLleuArgbArgbArgbArgbArgbGlnGlnuArgbGln 815
QY 544 ---CGTCTAAGAAAGTTGCAAGAACAAAGCGATTTGAAACAAAGAGACCT 597
Db 816 LeuGlnuArgbGlnGlnuLgluLleuArgbArgbArgbArgbArgbGlnGlnuArgbGln 835
QY 598 GCTAAGAAAGTTGCAAGAGCAAGCAAGAGTTTGAACAAAGAGCGCTATACGAA 657
Db 836 ---GlnGlnuArgbGlnGlnuLgluLleuArgbArgbArgbArgbArgbGlnGlnuArgbGln 853
QY 658 AAAAATTGTAAGAAAGAAAGCACTGAGATATATTATGACAGATTATATGCTGCT 717
Db 854 ArgbArgbGlnuArgbGlnuLgluLleuArgbArgbArgbArgbArgbGlnGlnuArgbGln 870
QY 718 TTGAATATACAGCTATAGACTTCATCAGAAATGAAAGTGGATATATATACCAAT 777
Db 871 GluGln-----GluLeuArgbArgbArgbArgbArgbGlnGlnuArgbGln 886
QY 778 CAATCTTCTTAACTTCAGACAAAGAGGAATATGATGAT----- 819
Db 887 ArgbGlnuLleuGlnuLgluLgluLuarGlnuArgbArgbArgbArgbArgbGlnu 906
QY 819 ----- 819
Db 907 GlnLeuArgbGlnuArgbGlnuLgluLleuArgbArgbArgbArgbGlnuLcybArgbGln 926
QY 820 TCCAAAGAAATATCTATATAGAAAAAAACAAATAGAGAA-----TCTATT 864

[illegible]

FR		02-OCT-2002; 2002US-0414971P.	
XX			
PA	(GETH) GENENTECH INC.		
XX			
PI	Mu TD, Zhang Z, Zhou Y;		
XX			
DR	WPI; 2004-347921/32.		
XX	N-PSDB; ACN37449.		
PT	New tumor-associated antigenic target polypeptides and nucleic acids,		
PT	useful in preparing a medicament for treating or detecting a		
PT	proliferative disorder, e.g. breast, lung, colorectal, ovarian or		
PT	prostate cancer or tumor.		
XX			
PS	Claim 12; SEQ ID NO 339; 7273pp; English.		
XX			
CC	The invention relates to human tumour-associated antigenic target (TAT)		
CC	polypeptides, and their related nucleic acids. The TAT polypeptides are		
CC	overexpressed in cancer tissues compared to normal tissues, and may thus		
CC	serve as effective targets for the diagnosis and treatment of cancer in		
CC	mammals. The invention also relates to nucleic acid and polypeptide		
CC	sequences at least 80% identical to the TAT nucleic acids and		
CC	polypeptides; expression vectors and host cells comprising a TAT nucleic		
CC	acid; an antibody specific for a TAT polypeptide; a peptide or organic		
CC	molecule which binds to a TAT polypeptide; fusion proteins comprising a		
CC	TAT polypeptide; and methods and compositions for the treatment or		
CC	diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,		
CC	antibodies, antagonists, binding molecules and compositions are useful		
CC	for diagnosing or treating a cell proliferative disorder associated with		
CC	increased TAT expression, particularly cancers such as breast cancer,		
CC	colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder		
CC	cancer, pancreatic cancer, cervical cancer, cancers of the central		
CC	nervous system, melanoma and leukaemia. TAT nucleic acids may further be		
CC	used as hybridisation probes, in chromosome and gene mapping, in		
CC	chromosome identification and in gene therapy. The present sequence		
XX	represents a TAT polypeptide of the invention		
SQ	Sequence 1564 AA:		
	Alignment Scores:		
	Pred. No.: 1.68e-19 Length: 1564		
	Score: 373.00 Matches: 133		
	Percent Similarity: 46.53% Conservative: 115		
	Best Local Similarity: 24.95% Mismatches: 185		
	Query Match: 15.00% Indels: 100		
	DB: 8 Gaps: 21		
	US-09-837-344-41 (1-1482) x AEW80133 (1-1584)		
OY	1 CAAGAACAACAAAGCGCATCTAGAACCAAGAGACGCTGTAAAGAAAAGTTGCAGAACAA	60	
DB	278 ArgTgTngIngtlnuArGaRgTgIngtIngtlnuArGaRgTgIngtlnuArGleuYbArGIn	29	
OY	61 CAAAGCGATTGTAAGCAAGATAGACTTGTCCTAAAGAAAAGTTACAGAGCAGCAAAGCGAT	12	
DB	298 Glu-----GltngIngtlnuArGleuGInglArGleuArYserGIngtIngtlnuArG	31	
OY	121 TTAGAACAGAGAGACTGCTCTAAAGAAAAGTTG---CAAGACAAACAAAGCGATCTAGAA	17	
DB	315 ArgTgTngIngtlnuArGaRgTgIngtIngtlnuLeuLtybArGIngtlnuYbArGleuGIn	33	
OY	178 CAAAGAGACGCTGTAAAGAAAAGTTGCCAAGAACCAAAAGC-----	21	
DB	335 GIngtlnuArGaRgTgIngtlnuArGleuYbArGIngtIngtlnuArGArGseGIngtlnuLeu	35	
OY	220 GATTTAGAACCAAGAGACGCTGTAAAGAAAAGTTGCCAAGAACCAAAAGCGATTAGAA	27	
DB	355 LysArTgTngIngtlnuArGaRgTgIngtlnuArGleuLtybArGIngtlnuGIngtlnuArGleuGIn	37	
OY	280 CAAGATGACTGCTTAAAGAAAAGTTACAA-----GAGCAGCAAAGCGATTTAGAACAA	33	
DB	375 Gln--ArgleuYbArGIngtlnuValGlnuArGleuGIngtlnuGIngtlnuArGArGIngtlnu	39	

[illegible][illegible]

[illegible]

```

Oy 1081 GAAATATGATGAAGCTTAGAGCAATTTAGCGATATGCGAATATGCAATTTCCAGAT 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 341 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1141 GAAAGAAAACATAGCAATTTATTAAGACTAGAGATTTGATAGAGAAAATGAATTTA 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 AspGlnIleIleGlyPheGlyGlnGlnIleuLeuAsn-----ProGlnGlnGlnThrLeu 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1201 GATGATTTAGAT 1212
    |||
Db 378 GlnAlaPheAsp 381

```

Search completed: December 7, 2005, 06:48:53
Job time : 172.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 7, 2005, 06:33:52 ; Search time 31.5 Seconds

(Without alignments)
7779.388 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 2486 1 CAAACAACACAAAGCCATCT.....AATATTATGAACATATA 1482

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+np.model -DEV=x1h
-O=cgnt2_1/USPTO.apool h/US09837344/runat.05122005.133407.29484/app.query.fasta_1.1671
-DB=Issued_Patents_AA -OPMT=faetan -SUFFIX=rai -MINMATCH=0.1 -LOOPT=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09837344@CGN 1.1.34@runat.05122005.133407.29484 -NCPU=6 -ICPU=3
-NO MMAP -IARGOUIRRY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued Patents_AA:
1: /cgnt2_6/ptodata/1/1aa/5/COMB.pep:*
2: /cgnt2_6/ptodata/1/1aa/6/COMB.pep:*
3: /cgnt2_6/ptodata/1/1aa/H/COMB.pep:*
4: /cgnt2_6/ptodata/1/1aa/PTUS/COMB.pep:*
5: /cgnt2_6/ptodata/1/1aa/RE/COMB.pep:*
6: /cgnt2_6/ptodata/1/1aa/backfill1est.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1549	62.3	316 2	US-08-098-327E-31 Sequence 31, Appl
2	1549	62.3	316 2	US-08-462-625-31 Sequence 31, Appl
3	603	24.3	117 2	US-08-098-337E-20 Sequence 20, Appl
4	603	24.3	117 2	US-08-462-625-20 Sequence 20, Appl
5	554	22.3	107 2	US-08-098-327E-19 Sequence 19, Appl
6	554	22.3	107 2	US-08-462-625-19 Sequence 19, Appl
7	376	15.1	1898 1	US-08-056-200-94 Sequence 94, Appl
8	376	15.1	1898 1	US-08-800-644-94 Sequence 94, Appl
9	376	15.1	1898 2	US-09-538-092-1280 Sequence 1280, Ap
10	359	14.4	1162 1	US-08-728-323A-2 Sequence 2, Appl
11	359	14.4	1162 1	US-09-298-568-2 Sequence 2, Appl
12	359	14.4	1162 2	US-09-410-399-2 Sequence 2, Appl

13	359	14.4	1162 2	US-09-894-273-2 Sequence 2, Appl
14	336	13.5	608 2	US-09-270-767-32937 Sequence 32937, A
15	336	13.5	608 2	US-09-270-767-48154 Sequence 48154, A
16	323.5	13.0	611 2	US-09-216-393B-81 Sequence 81, Appl
17	319	12.8	1972 2	US-08-875-435B-3 Sequence 3, Appl
18	317.5	12.8	1937 2	US-09-538-092-918 Sequence 918, App
19	316	12.7	1180 2	US-09-543-681A-6436 Sequence 6436, Ap
20	315.5	12.7	1564 2	US-10-144-198-2 Sequence 2, Appl
21	315.5	12.7	1564 2	US-10-144-198-4 Sequence 4, Appl
22	315	12.7	1960 2	US-09-538-092-1077 Sequence 1077, Ap
23	315	12.7	1960 2	US-09-849-016-10872 Sequence 10872, A
24	314	12.6	1976 2	US-09-538-092-1078 Sequence 1078, Ap
25	309	12.4	1972 2	US-08-875-435B-4 Sequence 4, Appl
26	306.5	12.3	1942 2	US-09-849-016-8155 Sequence 8155, Ap
27	305	12.3	1939 2	US-09-538-092-915 Sequence 915, App
28	305	12.3	1939 2	US-09-849-016-11104 Sequence 11104, A
29	304.5	12.2	1939 2	US-09-849-016-6925 Sequence 6925, Ap
30	303.5	12.2	1939 2	US-09-248-796A-18798 Sequence 18798, A
31	303	12.2	1939 2	US-09-310-187A-1 Sequence 1, Appl
32	303	12.2	1939 2	US-09-538-092-917 Sequence 917, App
33	302	12.1	1959 2	US-09-849-016-8134 Sequence 8134, Ap
34	301	12.1	1857 2	US-09-917-254-91 Sequence 91, Appl
35	301	12.1	1972 2	US-09-538-092-1084 Sequence 1084, Ap
36	301	12.1	1984 2	US-09-949-016-7111 Sequence 7111, Ap
37	301	12.1	1984 2	US-09-949-016-7112 Sequence 7112, Ap
38	301	12.1	1984 2	US-09-949-016-7113 Sequence 7113, Ap
39	299.5	12.0	1886 2	US-08-938-105-3 Sequence 3, Appl
40	299.5	12.0	1935 2	US-09-538-092-916 Sequence 916, App
41	299	12.0	1878 2	US-09-914-259-11 Sequence 11, Appl
42	298.5	12.0	1944 2	US-09-949-016-10929 Sequence 10929, A
43	298	12.0	785 2	US-10-164-595-80 Sequence 80, Appl
44	298	12.0	784 2	US-10-164-595-79 Sequence 79, Appl
45	298	12.0	843 2	US-10-164-595-54 Sequence 54, Appl

ALIGNMENTS

RESULT 1
US-08-098-327E-31
; Sequence 31, Application US/08098327E
; Patent No. 6270771
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; TITLE OF INVENTION: DRUGS, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098.327E
; FILING DATE: 24-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620

Db 221 LyeLyeS1uH8G1yAsp11eLeuA1aG1uAspLeuTyrG1yArgLeuG1u1LeProA 240

Qy 733 ATAGACTTCATCAGAAAATGAACTGATATATATATACCATCAATCTTCTTAACT 792

Pred. No.:	1.09e-117	Length:	316
Score:	1569.00	Matches:	314
Percent Similarity:	99.37%	Conservative:	0
Best Local Similarity:	99.37%	Mismatches:	2
Query Match:	62.31%	Indels:	0
DB:	2	Gaps:	0

[illegible]

Qy	730	GCATAGACTTCACCTAGAAAATGAAACGTCGATATTATATACACATCAATCTTCTTTA	789
Db	41	Alalleltlueuprosergluaenlguaiglytyrtyrlleprohlslnserleu	60
Qy	790	CCTGAGACAACAGAGGGAATAGAGATTCCAGGAAATATCTATATAGAAAAACA	849
Db	61	ProclnlspsnhrnglylaensserArgspserlylelelellelglyleThr	80
Qy	850	AATAGACATCTATTACACAAATGTTGAAGACGAAAGGATATCACTAAAGACATCTT	909
Db	81	AsnArglylserlelleThrThrAsnValgluglylARgArspIlehlslslyghlslu	100
Qy	910	GAAGAAAAGAAAGATGTTCAATTAACCGACAAACAAAAGATTAATCT	960
Db	101	GluglulyslysbaspGlyserllelypsrogluglnlysluaplysser	117
RESULT 5			
US-08-098-327E-19			
Sequence 19, Application US/08098327E			
Patent No. 6270771			
GENERAL INFORMATION:			
APPLICANT: GUERIN-MARCHAND, Claudine			
APPLICANT: DRUIHE, Pierre			
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE			
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE			
OF STIMULATING THE T LYMPHOCYTES			
NUMBER OF SEQUENCES: 46			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Burns, Doane, Swecker & Mathis			
STREET: P. O. Box 1404			
CITY: Alexandria			
STATE: Virginia			
COUNTRY: United States			
ZIP: 22313-1404			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/098.327E			
FILING DATE: 24-NOV-1993			
CLASSIFICATION: 424			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: FR 91 01286			
FILING DATE: 05-FEB-1991			
ATTORNEY/AGENT INFORMATION:			
NAME: McGowan, Malcolm K.			
REGISTRATION NUMBER: 39,300			
REFERENCE/DOCKET NUMBER: 010830-045			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (703) 836-6620			
TELEFAX: (703) 836-2021			
INFORMATION FOR SEQ ID NO: 19:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 107 amino acids			
TYPE: amino acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE: peptide			
PUBLICATION INFORMATION:			
DOCUMENT NUMBER: WO 92/13884			
PUBLICATION DATE: 20-AUG-1992			
US-08-098-327E-19			
Alignment Scores:			
Pred. No.:			
Score:	3.71e-37	Length:	107
Percent Similarity:	554.00%	Matches:	107
Best Local Similarity:	100.00%	Conservative:	0
Query Match:	22.28%	Mismatches:	0
DB:	2	Indels:	0
		Gaps:	0

US-09-837-344-41 (1-1482) x US-08-098-327E-19 (1-107)

QY 640 AGGAGGCTGATACGAAAAAATTGAGAAAAAGAACATGAGATATATTACCA 699
1 ArglyeAlaaphrlylsylsbaenleuGluArglylsylslylserlleuAla 20

QY 700 GAGATTATATGCTGTTTGAATAATCCAGCTTATGAACTTCCATAGAAAAATGAACT 759
21 GluAspLeuTyrglyArgleuGluileProAlalleuLeuProserGluAsnGluArg 40

QY 760 GATATATATACACATCAATCTTTTACCTCAGACACAGAGGAATAGTAGAGAT 819
41 GlyTyrrlyleProHsglnSerSerleuProGlnAspAsnArgGlyAsnSerArgAsp 60

QY 820 TCCAGGAAATATCTATATAGAAAAACAATAGAAATCTATTACAAATGTTGAA 879
61 SerlyseGlnleueserlleleGlnlyleThrAsnArgGlnserlleThrThrAsnValGln 80

QY 880 GAGCAAGGATATACATAAGACATCTTGAGAAAAAGAAAGATGCTTCAATAAACCA 939
81 GlyArgAlaaphrlylslyslyslsleuGlnGlnlylsAspGlySerllelyPro 100

QY 940 GAACAAAAGAGATAATCT 960
101 GluGlnlysgluAspLysSer 107

RESULT 6

US-08-462-625-19
Sequence 19, Application US/08462625
Patent No. 6319502

GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-19

Alignment Scores:
Pred. No.: 3,71e-37 Length: 107
Score: 554.00 Matches: 107
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.28% Indels: 0
DB: 2 Gaps: 0

US-09-837-344-41 (1-1482) x US-08-462-625-19 (1-107)

QY 640 AGGAGGCTGATACGAAAAAATTGAGAAAAAGAACATGAGATATATTACCA 699
1 ArglyeAlaaphrlylsylsbaenleuGluArglylsylslylserlleuAla 20

QY 700 GAGATTATATGCTGTTTGAATAATCCAGCTTATGAACTTCCATAGAAAAATGAACT 759
21 GluAspLeuTyrglyArgleuGluileProAlalleuLeuProserGluAsnGluArg 40

QY 760 GATATATATACACATCAATCTTTTACCTCAGACACAGAGGAATAGTAGAGAT 819
41 GlyTyrrlyleProHsglnSerSerleuProGlnAspAsnArgGlyAsnSerArgAsp 60

QY 820 TCCAGGAAATATCTATATAGAAAAACAATAGAAATCTATTACAAATGTTGAA 879
61 SerlyseGlnleueserlleleGlnlyleThrAsnArgGlnserlleThrThrAsnValGln 80

QY 880 GAGCAAGGATATACATAAGACATCTTGAGAAAAAGAAAGATGCTTCAATAAACCA 939
81 GlyArgAlaaphrlylslyslyslsleuGlnGlnlylsAspGlySerllelyPro 100

QY 940 GAACAAAAGAGATAATCT 960
101 GluGlnlysgluAspLysSer 107

RESULT 7

US-08-056-200-94
Sequence 94, Application US/08056200
Patent No. 5616500

GENERAL INFORMATION:
APPLICANT: Steinert, Peter M.
APPLICANT: Lee, Seung-Chul
APPLICANT: Kim, In-Gyu
APPLICANT: Chung, Soo-Il
APPLICANT: Park, Sang-Chul
TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
TITLE OF INVENTION: Methods of Using Same
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/056,200
FILING DATE: 30-APR-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Fredrick, Michael P.
REGISTRATION NUMBER: 36,799
REFERENCE/DOCKET NUMBER: NIH054.001A

```
TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1898 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-056-200-94

Alignment Scores:
Pred. No.:      1,44e-22      Length:      1898
Score:          376.00      Matches:      137
Percent Similarity: 47.3%      Conservative: 117
Best Local Similarity: 25.56%      Mismatches: 174
Query Match:    15.12%      Indels:      108
DB:              1          Gaps:      24

US-09-837-344-41 (1-1482) x US-08-056-200-94 (1-1898)

QY      1 CAAGAACAACAAGCGATCTAGAACAGAGACGTGCTTAAAGAAAAGTTGCAAGACAA 60
      ::::::::::::::::::::|
Db      522 ArgGluGlnGlnLeuArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 541
QY      61 CAAGCGATTTGACACAGATAGACTTGCTTAAAGAAAAGTTGCAAGACAGCAAGCGAT 120
      ::::::::::::::::::::|
Db      542 Glu-----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 558
QY      121 TTAGAACA---GAGAGACTTGCTTAAAGAAAAGTTGCAAGACAAACAAGCATCTAGAA 177
      |||||
Db      559 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 578
QY      178 CAAGAGACGCTGCTTAAAGAAAAGTTGCAAGACAAACAAGCGATTTGAA 219
      ::::::::::::::::::::|
Db      579 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 598
QY      220 GATTTGAACAAGAGACGTGCTTAAAGAAAAGTTGCAAGACAAACAAGCATTTGAA 279
      |||||
Db      599 LysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 618
QY      280 CAAGATAGACTTGCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCATTTGAA 339
      |||||
Db      619 Gln---ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 633
QY      340 CGTGTAAAGAAAAGTTG-----CAAGACAAACAAGCGATTTA--- 378
      |||||
Db      634 ArgArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 653
QY      379 -----GAACAAAGAGACGTGCTTAAAGAAAAGTTGCAAGACAAACAAGCATTTA 429
      |||||
Db      654 LysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 673
QY      430 GAACAAGAGAGACTTGCTTAAAGAAAAGTTGCAAGACAAACAAGCGATTTA 474
      ::::::::::::::::::::|
Db      674 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 693
QY      475 GATTTGAACAAGAGACGTGCTTAAAGAAAAGTTGCAAGACAAACAAGCGATTTA 525
      ::::::::::::::::::::|
Db      694 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 713
QY      526 -----GATTGGAACAAGAG----- 540
      |||||
Db      714 GluArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 733
QY      541 AGACGCTTAAAGAAAAGTTGCAAGACAAACAAGCGATTTA 579
      |||||
Db      734 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 753
QY      580 TTAGAACAAGAGACGTGCTTAAAGAAAAGTTGCAAGACAAACAAGCGATTTA 615
      |||||
Db      754 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 773
```

```
QY      616 GAGCAGCAAGAGATTTAGAACAAAGAGCGCTGATACG---AAAAAAATTTAGAAAGA 672
      |||||
Db      774 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 793
QY      673 AAAAAGAACATGAGATATATTTAGCAGAGATTTATATGTCGT----- 717
      ::::::::::::::::::::|
Db      794 LysSerGlnArgGly-----ArgGlnArgLeuSerAlaArgProLeuArgGln 810
QY      718 -----TTAGAAATACACGCTATAGAACTTCATCAGAAATGAACTGATATAT 768
      ::::::::::::::::::::|
Db      811 GlnArgGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 828
QY      769 ATACCATCATCATCTTCTTACTCAGGACAAACAGCGGAATAGAGATTTCCAGGAA 828
      ::|||
Db      829 LeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 848
QY      829 ATATCTATATATGAAAAAAACAATAGA-----GAATCTATTACA 867
      ::|||
Db      849 LeuGlnPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 868
QY      868 ACAATGTTGAAGA---CGAAGGATATACATTAAGACATCTTGGAAGAAAAGAAAGAT 924
      ::|||
Db      869 GluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 888
QY      925 GGTTCATTAACCAAGCAAGAAAAGAGATTAATCTGCTGACATTAACAATCATCATTA 984
      ::|||
Db      889 GlnLysTrpArgTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 906
QY      985 GAGACAGTAAATATTTCTGATGTTAATGATTTTCAATTAAGTAACTATAGAGATCAATA 1044
      ::|||
Db      907 -----TyrAlaLysProAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 916
QY      1045 AGTGTGAATATGACGATTCATTAATAGATTAAGAAAGAAATGATGAAGACTTAGACAA 1104
      |||||
Db      917 ArgLysGln-----GlnGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 934
QY      1105 TTTAAGCTATTTGCAATATGACATTTCCAAAGTGAAGAAAACATAGAAATTTATATA 1164
      ::|||
Db      935 -----ArgGlnLysArgArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 948
QY      1165 GAACTAGAAGATTTGATAGAGAAAATGAAAATTTAGATTTAGATGAAGAAATAGAA 1224
      |||||
Db      949 GluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 966
QY      1225 AAATCATCAGAAAGATTTCTGAGACAAAATTAAGAAAAGAAAGAAATATGAAAACAA 1284
      ::|||
Db      967 LysArgArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 986
QY      1285 AAGGATATATATTT---AAACCAATGATTAAGATTTGATGATGACCATTTAAAAAA 1341
      ::|||
Db      987 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1006
QY      1342 TATTAATAATGATTAAGCAGTTTAAAGGAAAAGAAAATTCATATAAA 1389
      ::|||
Db      1007 TyrArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1022

RESULT 8
US-08-644-94
; Sequence 94, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichobyalin and Transglutaminase-3 and
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
```

STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,644
FILING DATE: 14-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/056,200
FILING DATE: 30-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fedrick, Michael F.
REGISTRATION NUMBER: 36,799
REFERENCE/DOCKET NUMBER: NIH054,001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (714) 760-0404
TELEFAX: (714) 760-9502
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 1898 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-800-644-94

Alignment Scores:

Pred. No.:	1,44e-22	Length:	1898
Score:	376.00	Matches:	137
Percent Similarity:	47.39%	Conservative:	117
Best Local Similarity:	25.56%	Mismatches:	174
Query Match:	15.12%	Indels:	108
DB:	1	Gaps:	24

US-09-837-344-41 (1-1482) x US-08-800-644-94 (1-1898)

QY 1 CAAGAACAAAGCGATCTAGAACAAAGAGAGCGTCTAAAGAAAGTTGCAAGAACAA 60
DB 522 ArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIu 541
QY 61 CAAAGCGATTTAGAACAAAGATGACTTCTAAAGAAAGTTTCAAGAGCAGCAAGCGAT 120
DB 542 GIu-----GIuGIuGIuArgIuGIuInIeuArgIuArgserGIuGIuInIeuArg 558
QY 121 TTAGAACAA---GAGAGCTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAA 177
DB 559 ArgIuGIuInIeuArgIuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArg 578
QY 178 CAAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAA 219
DB 579 GIuGIuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArg 598
QY 220 GATTAGAACAAAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAA 279
DB 599 LySerArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArg 618
QY 280 CAAGATTAAGCTTCTAAAGAAAGTTTCAAGAGCAGCAAGCGATTTGACCAAGAGGA 339
DB 619 GIu---ArgIeuLySerGIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArg 633
QY 340 CGTGTAAAGAAAGTTG-----CAAGAACAAAGCGATTTA--- 378
DB 634 ArgArgArgGIuArgIeuLySerGIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArg 653
QY 379 -----GAACAAAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 429
DB 654 LySerArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArg 673
QY 430 GAACAAAGAGAGCGTCTAAAGAAAGTTGCAAGAACAA-----CAAGC 474

DB 674 ArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIu 693
QY 475 GATTAGAACAAAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 525
DB 694 GIuInIeuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIu 713
QY 526 -----GATTAGAACAAAG----- 540
DB 714 GIuArgIeuLySerArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIu 723
QY 541 AGAGCTCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 579
DB 734 ArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIu 753
QY 580 TTAGAACAAAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 615
DB 754 GIuGIuGIuLySerArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIu 773
QY 616 GAGCAGCAAGAGATTTAGAACAAAGAGCGATGATACG---AAAAAAATTTAGAACAA 672
DB 774 ArgIuGIuInIeuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArg 793
QY 673 AAAAAGAACATGAGATATATTTAGCAGAGATTTATATGCTGCT----- 717
DB 794 LySerArgIuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArg 810
QY 718 -----TTAGAACAAAGCGATTTAGAACCTTCATCAGAACAAATGAGATATATAT 768
DB 811 GIuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIu 828
QY 769 ATACCAATCAATCTTCTTACCTCAGACCAAGAGGAAATAGTAGATTTCCAGAACAA 828
DB 829 LeuProGIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIu 848
QY 829 ATATCTAATATAGAAAAAACAATAG-----GAATCTATTTACA 867
DB 849 LeuGIuInIeuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIu 868
QY 868 ACAAAATTTAGAGAG---CGAAGGATATATCAATTAAGACATCTTGAAGAAAGAAAGAT 924
DB 869 GIu 888
QY 925 GGTTCATTAACCAAGAACAAAGAGATTAATCTCTGACATCAAAATCATACATTA 984
DB 889 GIuLySerArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArgIuGIuInIeuArg 906
QY 985 GAGCAGTAATATATTTCTGATGTTAATGATTTTCAATTAAGTAGATGAGATGAATA 1044
DB 907 -----TyrAlaLySerProAlaLeuGIuGIuInIeu 916
QY 1045 AGTGTGAATATGACGATTCATTAATGATGAGAGAGAGATGATGAAGCTTGAAGAGAA 1104
DB 917 ArgLySerGIu-----GIuGIuInIeuGIuGIuInIeuGIuGIuInIeuGIuGIuInIeuArg 934
QY 1105 TTTAAGCTATGTGCAATATGACAAATTTCCAGAGTAAGAAACATAGGAATTTATATA 1164
DB 935 -----ArgGIuLySerArgIuGIuInIeuGIuGIuInIeuGIuGIuInIeuArg 948
QY 1165 GAATAGAGATTTGATAGAGAAATGAAATTTAGATGATTTAGATGAGAGAAATAGAA 1224
DB 949 GIu 966
QY 1225 AATATCATCAGAGAAATTTCTGAAGAAAGAAATTTAAAGAAAGAAATATGAAAGAAACA 1284
DB 967 LySerArgIuArgIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 986
QY 1285 AAGATATATATTT---AAACCAATGATTAAGTTTGTATGTAGCATATTTAAAAAA 1341
DB 987 GIu 1006
QY 1342 TATTAATATGATAGCAGTTAATAGAAAGAAATTTCAATATAA 1389

RESULT 10
US-08-728-323A-2

Sequence 2, Application US/08728323A

Patent No. 5948676

GENERAL INFORMATION:

APPLICANT: Chang, Yuan

APPLICANT: Bohenzky, Roy A.

APPLICANT: Russo, James J.

APPLICANT: Edelman, Isidore S.

TITLE OF INVENTION: Immediate Early Protein From Kaposi's

TITLE OF INVENTION: Sarcoma-Associated Herpesvirus DNA

TITLE OF INVENTION: Encoding Same And Uses Thereof

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/728,323A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 0575/52268/JPM/MS/SKS

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-278-0400

TELEFAX: 212-591-0525

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1162 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-728-323A-2

Alignment Scores:

Pred. No.: 3,26-21 Length: 1162

Score: 359.00 Matches: 94

Percent Similarity: 49.44% Conservative: 127

Best Local Similarity: 21.03% Mismatches: 110

Query Match: 14.44% Indels: 116

Gaps: 12

US-09-837-344-41 (1-1482) x US-08-728-323A-2 (1-1162)

4 GAACAACAAGCGACTTAGAACAGAGAGCGTCTTAAGAAAGTTGCAAGAACAA 63
|||||
604 G|U|G|I|n|G|I|n|A|e|p|---G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n| 622
64 AGCGATTAGAAACAAGATAGACTTCTTAAGAAAGTTGCAAA-----GAGCAGCAAGC 117
:::|
623 A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n| 642
QY 118 GATTAGAAACAAGAGACTTCTTAAGAAAGTTGCAAGAAACAACAAGCGATCTGAA 177
|||
Db 643 A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n| 662
QY 178 CAAGAGAGCGTCTTAAGAAAGTTGCAAGAAACAACAAGCGATTTAGAACAAAGAGA 237
|||
Db 663 A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n| 682
QY 238 CGTGTAAAGAAAGTTGCAAGAAACAACAAGCGATTTAGAAACAAGATAGACTTCTTAA 297

Db 683 A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n| 702
QY 298 GAAAAATTACAAAGCGCAAGACGATTTAGAACAAAGAGAGCGTCTTAAGAAAGTTG 357
:::|
Db 703 G|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n| 722
QY 358 CAAGAAACAAGCGATTTAGAAACAAGAGAGAGCGTCTTAAGAAAGTTGCAAGAAACA 417
|||
Db 723 G|I|n|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n|A|e|p|G|U|G|I|n|G|I|n| 742
QY 418 CAAGCGATTTAGAAACAAGAGAGCTTCTTAAGAAAGTTGCAAGAAACAACAAGCGAT 477
|||
Db 743 G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 758
QY 478 TTAGAACAAAGAGAGCGTCTTAAGAAAGTTGCAAGAAACAACAAGCGATTTAGAACAA 537
|||
Db 759 G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 778
QY 538 GAGAGAGCGTCTTAAGAAAGTTGCAAGAAACAACAAGCGATTTAGAACAAAGAGAGAG 597
:::|
Db 779 G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 793
QY 598 GCTTAAGAAAGTTGCAAGAGAGCAAGAGATTTAGAACAAAGAGAGCGTGTATACGAA 657
:::|
Db 794 ---G|U|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 808
QY 658 AAAAAATTAGAAAGAAAGAGACATGAGATATATTACAGAGAGATTTATATGTCGT 717
:::|
Db 809 G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 817
QY 718 TTAGAAATACAGGTATAGAACTTCATCAGAAATGAGTGTATATTATACCAT 777
|
Db 817 ----- 817
QY 778 CAATCTTCTTACCTCAGACAAACAGAGGAATAGAGATTCAGAGAAATATCTATA 837
|
Db 817 ----- 817
QY 838 ATAGAAAAACAATAGAGAAATCTATTACAAACAATGTTGAGAGCAGAGGATATACAT 897
:::|
Db 818 L|e|u|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 824
QY 898 AAAGACATCTTGAGAAAGAAAGAGATGTTCAATTAACAGACAAAGAAAGATATA 957
|||
Db 825 -----L|e|u|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 837
QY 958 TCTGCTGACATACAAATCATCATTAAGAGACAGTAATATTCTGATTTAATGATTTT 1017
|
Db 838 ----- 838
QY 1018 CAATTAAGTATAGATAGAGATGAATAAGTCTGAATATAGACGATTCATTAATATGATGAA 1077
:::|
Db 843 G|I|U|A|G|I|U|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 856
QY 1078 GAAGAAGATGTAAGACTTAGACGAATTTAAGCTTATGCGAATATGCAATTTCCAA 1137
|||
Db 857 G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 876
QY 1138 GATGAGAAAAATATGAATTTATTAAGAACTTAAGATTTGATAGAGAAATATGAA 1194
|||
Db 877 G|I|U|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 890
QY 1195 AATTAGATGATTTAGATGAGAAATATGAATAATATCATCAGAAAGATTTATGAGAA 1254
|||
Db 891 G|I|U|e|U|G|I|U|A|G|I|U|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 910
QY 1255 ATAAAAAAGAAATATGAAAAACAAGATTAATATTTAAACCAATGATATA 1314
:::|
Db 911 G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n|G|I|n| 930
QY 1315 AGTTGTATGATGAGCATATT 1335
:::|

Db 931 GluThrValGluGluProIle 937

RESULT 11

US-09-298-568-2

Sequence 2, Application US/09298568

Patent No. 6322792

GENERAL INFORMATION:

APPLICANT: Kieff, Elliott D.

APPLICANT: Ballester, Mary E.

APPLICANT: Kaye, Kenneth M.

TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO

TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE

FILE REFERENCE: 16412-10001R

CURRENT APPLICATION NUMBER: US/09/298,568

EARLIER FILING DATE: 1999-04-21

NUMBER OF SEQ ID NOS: 3

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 1162

TYPE: PRT

ORGANISM: Kaposi's sarcoma-associated herpesvirus

US-09-298-568-2

Alignment Scores:

Pred. No.:	3,2e-21	Length:	1162
Score:	359.00	Matches:	94
Percent Similarity:	49.44%	Conservative:	127
Best Local Similarity:	21.03%	Mismatches:	110
Query Match:	14.44%	Indels:	116
Db:	2	Gaps:	12

US-09-837-344-41 (1-1482) x US-09-298-568-2 (1-1162)

QY 4 GAACAACAAGCGATCTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAACAA 63

Db 604 GluGlnGlnGlnAsp---GluGlnGlnGlnAspGluGlnGlnGlnGlnGlnGln 622

QY 64 AGCGATTTCACACATAGACTCTCTAAAGAAAGTTGCAAGAACAA 117

Db 623 AspGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 642

QY 118 GATTTCAGACAGAGAGACTCTCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAA 177

Db 643 AspGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 662

QY 178 CAAGAGAGCGTCTTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGA 237

Db 663 AspGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 682

QY 238 CGTCTTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGTCTTAA 297

Db 683 AspGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 702

QY 298 GAAAGATTTCAGAGAGAGCGATTTAGAACAGAGAGCGTCTTAAGAAAGTTG 357

Db 703 GlnAspGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 722

QY 358 CAAGAACAAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAGTTGCAAGAACAA 417

Db 723 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 742

QY 418 CAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAGTTGCAAGAACAAAGCGAT 477

Db 743 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 758

QY 478 TTGACACAGAGAGCGTCTTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAA 537

Db 759 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 778

QY 538 GAGAGAGCGTCTTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAG 597

Db 538 GAGAGAGCGTCTTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAG 597

Db 779 Gln-----GluGlnGlnLeuGluGlnGlnGlnGlnGlnGlnGlnGln----- 793

QY 598 GCTAAAGAAAGTTGCAAGAGAGCGAACAAAGATTTAGAACAAAGAGCGTATACGAA 657

Db 794 ---GluGlnGlnLeuGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln-----Glu 808

QY 658 AAAAATTTAGAAAGAAAGAAAGAAAGCGATATATTAGCAGAGATTTATATGTCGT 717

Db 809 GlnGlnLeuGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 817

QY 718 TTAGAAATACAGCATATGAACTTCACAGAAATAGAACGCGATATTATATACACAT 777

Db 817 ----- 817

QY 778 CAATCTCTTACCTCAGACAGACAGAGAGATAGAGATTCAGAGAAATATCTATA 837

Db 817 ----- 817

QY 838 ATGAAAAACAAATAGAAATCTATTCAACAAATGTTGAAGCAGAAAGGATATACAT 897

Db 818 LeuGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 824

QY 898 AAAGCACTCTTGAAGAAAGAAAGATGTTCAATAAACAGACAAAGAGATATAA 957

Db 825 -----LeuGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 837

QY 958 TCTGCTGACATCAAAATCATATATAGAGACAGTAAATATTTCTGATGTTAATGATTT 1017

Db 838 -----GluLeuGlnGlnGln 842

QY 1018 CAATATAGATATAGAGATGAAATAGCTGTAATAGAGATTCATTAATAGATGAA 1077

Db 843 GluValGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 856

QY 1078 CAAGAAAGATGATGAGATCTTGAAGATTTAAGCTATTTGCAATATGACATTTCCAA 1137

Db 857 GluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 876

QY 1138 GATGAAGAAACATAGCAATTTATTAAGAAATCTAGAAAGTTGATAGAGAAATGAA 1194

Db 877 GluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 890

QY 1195 AATTGATGATTTAGATGAGAAATAGAAATATCTAGAAATTTATCTGAGAAAGAA 1254

Db 891 GluLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 910

QY 1255 ATTAAGAAAGAAAGAAATATGAGAAAGAAAGATTAATTTAAACCAATGATATA 1314

Db 911 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 930

QY 1315 AGTTGATGATGACATATTT 1335

Db 931 GluThrValGluGluProIle 937

RESULT 12

US-09-410-399-2

Sequence 2, Application US/09410399

Patent No. 6482587

GENERAL INFORMATION:

APPLICANT: Robertson, Erle S.

APPLICANT: Cotter, Murray A.

TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA

TITLE OF INVENTION: to Genomic Host DNA

FILE REFERENCE: UM-03778

CURRENT APPLICATION NUMBER: US/09/410,399

EARLIER FILING DATE: 1999-10-01

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 1162

TYPE: PRT

ORGANISM: Kaposi's sarcoma-associated herpesvirus

US-09-410-399-2

Alignment Scores:
Pred. No.: 3.2e-21 Length: 1162
Score: 359.00 Matches: 94
Percent Similarity: 49.44% Conservative: 127
Best Local Similarity: 21.03% Mismatches: 110
Query Match: 14.44% Indels: 116
DB: 2 Gaps: 12

US-09-837-344-41 (1-1482) x US-09-410-399-2 (1-1162)

4 GAACAAACGCGATTCTAGAACAGAGAGCGCTCTAAAGAAAAGTTGCAAGAACAA 63
|||||
604 GluInglInglInp---GluInglInglInpGluInglInglInpGluInglInglIn 622
64 AGCGATTTAGAACAGATAGACTTCTAAAGAAAAGTTACAA-----GAGCGAACAA 117
::: |||||
623 AspIuInglInglInpGluInglInglInpGluInglInglInpGluInglInglIn 642
118 GATTAGAACAGAGAGACTTCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAA 177
||| :|||
643 AspIuInglInglInpGluInglInglInpGluInglInglInpGluInglInglIn 662
178 CAAGAGAGCGCTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGA 237
||| :|||
663 AspIuInglInglInpGluInglInglInpGluInglInglInpGluInglInglIn 682
238 CGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATAGACTTCTAA 297
:|||||
683 AspIuInglInglInpGluInglInglInpGluInglInglInpGluInglInglIn 702
298 GAAAAGTTACAGAGAGCGAAAGCGATTTAGAACAGAGAGAGCGCTCTAAAGAAAAGTTG 357
::: |||||
703 GluInpGluInglInglInpGluInglInglInpGluInglInglInpGluInglIn 722
358 CAAGAACAGAGAGCGATTTAGAACAGAGAGCGCTCTAAAGAAAAGTTGCAAGAACAA 417
||| :|||
723 GluInglInglInpGluInglInglInpGluInglInglInpGluInglInglIn 742
418 CAAGCGATTTAGAACAGAGAGACTTCTAAAGAAAAGTTGCAAGAACAAAGCGAT 477
||| :|||
743 GluInpGluInglInglInpGluInglInglInpGluInglInglInpGluInglIn 758
478 TTAGAACAGAGAGAGCGCTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAA 537
||| :|||
759 GluInglInglInpGluInglInglInpGluInglInglInpGluInglInglInp 778
538 GAGAGAGCGCTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGT 597
::: |||||
779 GluInglInglInpGluInglInglInpGluInglInglInpGluInglInglInp 793
598 GCTAAAGAAAAGTTGCAAGAGAGCGAAAGATTTAGAACAAAGAGCGTATACGAA 657
:|||||
794 ---GluInglInglInpGluInglInglInpGluInglInglInpGluInglInglInp 808
658 AAAAATTAGAACAGAAAAGAGAACTAGAGATATATTAGAGAGAGATTTATGTCGT 717
::: |||||
809 GluInglInglInpGluInglInglInpGluInglInglInpGluInglInglInp 817
718 TTAGAAATACAGCTATAGAACTTCATCAGAAAAATGAAGTATATTATACCAT 777
817 ----- 817
778 CAATCTCTTACTACGAGCAACAGAGGAATAGTAGATTCCAGAAATATCTATA 837
817 ----- 817
838 ATAGAAAAACAAATAGAAATCTATTAACAAATGTTGAAGAGAGGATATACAT 897
:|||||
818 LeuInglInglInglInp----- 824
898 AAAGACATCTTGGAAGAAAAGAGATGTTCAATAAACAGAACAAAGGATCTAGAA 957
||| :|||

825 -----LeuInglInglInglInglInglInglInglInglInglInglIn----- 837
QY 958 TCTGCTGACATACAAATATCATATTAGAGACAGCTAAATTTCTGATGTTAATGATTTT 1017
DB 838 -----GluLeuInglInglIn 842
1018 CAATATAGTAGTAGAGATGAGTAATAGTGTGAATATAGCGATTCATTAATAGATGAA 1077
QY 843 GluValGluInglInglInglInglInglInglInglInglInglInglIn-----GluIn 856
DB 1078 GAAGAAAGTAGTAGAGACTTACAGCAATTTAACCTTATGTCATATGCAATATGCAATTTCCAA 1137
QY 857 GluInglInglInglInglInglInglInglInglInglInglInglInglInglIn 876
DB 1138 GATGAGAAAAACATAGAAATTTATAAAGAACTGAAGATTTGATAGAGAAAATGAA--- 1194
QY 877 GluInglInglIn-----GluLeuInglInglInglInglInglInglInglIn 890
DB 1195 AATTAGATGATTAGTAGAGAAATAGAAAATCATCAGAGAAATTTATCTGAAGAAAAA 1254
QY 891 GluLeuInglInglInglInglInglInglInglInglInglInglInglInglInglIn 910
DB 1255 ATAAAAAAGGAAAATATAGAAAAAACAAGATTAATTTAAACCAAAATGATAAA 1314
QY 911 GluInglInglInglInglInglInglInglInglInglInglInglInglInglInglIn 930
DB 1315 AGTTGTATGATGACATATT 1335
QY 931 GluThrValGluInglInpProle 937

RESULT 13
US-09-894-273-2
Sequence 2, Application US/09894273
Patent No. 6756203
GENERAL INFORMATION:
APPLICANT: Kieff, Elliott D.
APPLICANT: Ballstaer, Mary E.
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
FILE REFERENCE: 16412-10001R
CURRENT APPLICATION NUMBER: US/09/894,273
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/109,422
PRIOR FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1162
TYPE: FRT
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-894-273-2

Alignment Scores:
Pred. No.: 3.2e-21 Length: 1162
Score: 359.00 Matches: 94
Percent Similarity: 49.44% Conservative: 127
Best Local Similarity: 21.03% Mismatches: 110
Query Match: 14.44% Indels: 116
DB: 2 Gaps: 12

US-09-837-344-41 (1-1482) x US-09-894-273-2 (1-1162)

4 GAACAAACGCGATTCTAGAACAGAGAGCGCTCTAAAGAAAAGTTGCAAGAACAA 63
|||||
604 GluInglInglInp---GluInglInglInpGluInglInglInpGluInglInglIn 622
64 AGCGATTTAGAACAGATAGACTTCTAAAGAAAAGTTACAA-----GAGCGAACAA 117
::: |||||
623 AspIuInglInglInpGluInglInglInpGluInglInglInpGluInglInglIn 642
118 GATTAGAACAGAGAGACTTCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAA 177
||| :|||

```
Db      643  AepglunglInglInaepglunglInglInaepglunglInglInaepglunglInglIn 662
Qy      178  CAAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTGTGAACAAGAGAGA 237
Db      663  AepglunglInglInaepglunglInglInaepglunglInglInaepglunglInglIn 682
Qy      238  CGTGTAAAGAAAAGTTGCAAGAACAAAGCGATTGTGAACAAGATGACTTGTAA 297
Db      683  AepglunglInglInaepglunglInglInaepglunglInglInaepglunglInglIn 702
Qy      298  GAAAAGTTACAAAGACGCAAAAGCGATTGTGAACAAGAGACGCTGCTAAAGAAAGTTG 357
Db      703  GlnaepglunglInglInaepglunglInglInaepglunglInglInaepglunglIn 722
Qy      358  CAAGAACAAACAAGCGATTGTGAACAAGAGACGCTGCTAAAGAAAGTTGCAAGAACAA 417
Db      723  GlnglInglInglInaepglunglInglInaepglunglInglInaepglunglInglIn 742
Qy      418  CAAGCGATTGTGAACAAGAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGAT 477
Db      743  GlnglInaepglunglInglIn-----GlnaepglunglInglInglInglInglIn 758
Qy      478  TTGAACAAGAGAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTGTGAACAA 537
Db      759  GlnglInglInglInglInglInglInglInglInglInglInglInglInglInaep 778
Qy      538  GAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTGTGAACAAGAGACGT 597
Db      779  Gln-----GlnglInglInleuInglInglInglInglInleuInglInglIn----- 793
Qy      598  GCTAAAGAAAAGTTGCAAGAGCAGCAAAAGATTTTGAACAAAGAAAGCGTATACGAA 657
Db      794  ---GlnglInglInleuInglInglInglInglInglInglInleuInglInglIn-----Gln 808
Qy      658  AAAAATTGAAGAAAAGAAAAGAACATGAGATATATATGACAGAGATTATATGCTGCT 717
Db      809  GlnglInleuInglInglInglInglIn----- 817
Qy      718  TTGAATAATCCAGCTATAGAACTTCATCAGAAAATGAACGTGATTTATATACCAAT 777
Db      817  ----- 817
Qy      778  CAATCTTTTACCTCAGGCAACAGAGGGAATAGTAGAGATTCCAAGAAAATATCTATA 837
Db      817  ----- 817
Qy      838  ATAGAAAAAACAAATAGAGATCTATTACAACAAATGTTGAAGACGAAAGGATATATCAT 897
Db      818  leuInglInglInglInglIn----- 824
Qy      898  AAAGCATCTTGAAGAAAAGAAAGATGCTTCAATTAACCAAGAACAAAGAAAGATTA 957
Db      825  -----leuInglInglInglInglInleuInglInglInglInglIn----- 837
Qy      958  TCTGCTGACATACAAATCATATCATTAGAGACAGTAATATTTCTGATGTTAATGATTT 1017
Db      838  -----GlnleuInglInglIn 842
Qy      1018  CAAATAGTAAGTAGAGATGAATAAGTCTGAATATGACGATTCATTAATAGATGA 1077
Db      843  GlnValaGlnglInglInglInValaGlnglIn-----GlnIn 856
Qy      1078  GAAGAAGATGATGAAGACTTAGACGAATTTAAGCTTATGCAATATGCAATTTCCAA 1137
Db      857  GlnglInglInglInglInleuInglInglInValaGlnglInglInglInglInglIn 876
Qy      1138  GATGAAGAAAACATAGGAATTTAATAAGAACTGAAGATTTGATAGAAAATGA--- 1194
Db      877  GlnleuInglIn-----GlnleuInglInValaGlnleuInglInglInglIn 890
Qy      1195  AATTAGATGATTTAGATGAAGAAATATATCATCAGAGAAATATCTGAAGAAAA 1254
Db      891  GlnleuInglInValaGlnleuInglInglInglInleuInglInValaGlnleuInglIn 910
```

```
Qy      1255  ATAAAAAGAAAGAAATATGAAAAAACAAAGATTAATTTTAAACCAATGATATA 1314
Db      911  GlnglInglInleuInglInglInValaGlnleuInglInglInglInglInValaGlnleuInglInglIn 930
Qy      1315  AGTTTATGATGACCATATT 1335
Db      931  GlnThrValGlnleuInleuProfile 937

RESULT 14
US-09-270-767-32937
; Sequence 32937, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 32937
; LENGTH: 608
; TYPE: PRN
; ORGANISM: Drosophila melanogaster
US-09-270-767-32937

Alignment Scores:
Pred. No.: 2,146-19 Length: 608
Score: 336.00 Matches: 70
Percent Similarity: 64.94% Conservative: 80
Best Local Similarity: 30.30% Mismatches: 73
Query Match: 13.52% Indels: 8
DB: 2 Gaps: 4

US-09-837-344-41 (1-1482) x US-09-270-767-32937 (1-608)
Qy      1  CAAGAACAAACAAGCGATCTAGAACACAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAA 60
Db      374  LysGlnGlnGlnGlnLysArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 393
Qy      61  CAAGCGATTGTGAACAAGATGACTTGTAAAGAAAAGTTTGAACAAAGCAATCTAGAACAA 120
Db      394  GlnArgArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 413
Qy      121  TTGAACAAGAGACCTGCTTAAAGAAAAGTTGCAAGAACAAAGCAATCTAGAACAA 180
Db      414  LysGlnLysArgArgArgArgArgLysArgArgArgArgLysArgLysArgLysGln 433
Qy      181  GAGAGCGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGT 240
Db      434  GlnArgLysGlnGlnGlnGlnArgArgLysArgLysArgLysArgLysArgLysArgLys 452
Qy      241  GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATGACTTGTAAAGAA 300
Db      453  AlaGlnGlnLysArgArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 472
Qy      301  AAGTTACAAAGACAGCAAGCGATTGTGAACAAGAGACGCTGCTTAAAGAAAGTTGCA 360
Db      473  LysArgArgArgArgGln-----GlnGlnLysArgLysGlnGlnGlnGlnGlnGlnGln 489
Qy      361  GAACAAACAAGCGATTGTGAACAAGAGACGCTGCT---AAAGAAAAGTTGCAAGAACAA 417
Db      490  GlnGlnGlnLysArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 509
Qy      418  CAAGCGATTGTGAACAAGAGACGCTGCTTAAAGAAAAGTTGCAAGAACAAAGCGAT 477
Db      510  GlnLysArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 529
Qy      478  TTGAACAAGAGACGCTGCTTAAAGAAAAGTTGCAAGAACAAAGCAATCTAGAACAA 537
Db      530  LysGlnGlnGlnLysArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 549
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: December 7, 2005, 06:55:32 ; Search time 130 Seconds
(without alignments)

9526.509 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 2486
Sequence: 1 CAAGAACACAAAGCCGATCT.....AATATTTCATGAACATATA 1482

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ nzp.model -DEV=xjh
-O=/cgnt2_1/USPTO.spool/h/US09837344/runac.05122005.133409.29551/app.query.fasta_1.1671
-DB=published Applications AA Main -OPMT=faster -SUFFIX=rapbm -MINMATCH=0.1
-LOOPC=0 -LOGEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62
-TRANS=unman40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09837344@CGN 1 1 174 @runac.05122005.133409.29551
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:

1: /cgnt2_6/ptodata/1/pubpaa/US07_PUBSCOMB.pep:*
2: /cgnt2_6/ptodata/1/pubpaa/US08_PUBSCOMB.pep:*
3: /cgnt2_6/ptodata/1/pubpaa/US09_PUBSCOMB.pep:*
4: /cgnt2_6/ptodata/1/pubpaa/US10A_PUBSCOMB.pep:*
5: /cgnt2_6/ptodata/1/pubpaa/US10B_PUBSCOMB.pep:*
6: /cgnt2_6/ptodata/1/pubpaa/US11_PUBSCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2406.5	96.8	1909	US-10-732-923-3341	Sequence 3341, Ap
2	1549	62.3	316	US-09-837-344-31	Sequence 31, Appl
3	1549	62.3	316	US-09-900-963-31	Sequence 31, Appl
4	603	24.3	117	US-09-837-344-20	Sequence 20, Appl
5	603	24.3	117	US-09-900-963-20	Sequence 20, Appl
6	554	22.3	107	US-09-837-344-19	Sequence 19, Appl
7	554	22.3	107	US-09-900-963-19	Sequence 19, Appl
8	381.5	15.3	1003	US-10-732-923-3319	Sequence 3319, Ap
9	380.5	15.3	1407	US-10-732-923-3359	Sequence 3359, Ap
10	380.5	15.3	1549	US-10-732-923-3358	Sequence 3358, Ap
11	376.5	14.3	419	US-10-450-763-47660	Sequence 47660, A
12	374	15.0	1439	US-10-754-079-157	Sequence 157, App

13	360	14.5	350	4	US-10-101-487-58	Sequence 58, Appl
14	360	14.5	350	5	US-10-939-988-58	Sequence 58, Appl
15	359	14.4	1152	3	US-09-894-273-2	Sequence 2, Appl1
16	359	14.4	1152	4	US-10-294-804-2	Sequence 2, Appl1
17	359	14.4	1152	5	US-10-194-046-2	Sequence 2, Appl1
18	337	13.6	1527	5	US-10-732-923-3354	Sequence 3354, Ap
19	331	13.3	2274	6	US-11-097-143-2763	Sequence 2763, Ap
20	323.5	13.0	611	3	US-09-321-956-81	Sequence 81, Appl
21	323.5	13.0	611	4	US-10-321-956-81	Sequence 81, Appl
22	320	12.9	658	4	US-11-097-143-23688	Sequence 23688, A
23	320	12.9	800	6	US-11-097-143-41169	Sequence 41169, A
24	318.5	12.8	1961	4	US-10-028-248A-103	Sequence 103, App
25	318.5	12.8	1961	4	US-10-107-782-103	Sequence 103, App
26	318	12.8	554	6	US-11-097-143-5154	Sequence 5154, Ap
27	317.5	12.8	1937	5	US-10-723-860-58	Sequence 58, Appl
28	317.5	12.8	1937	5	US-10-723-860-106	Sequence 106, App
29	315.5	12.7	1170	4	US-10-341-534-95	Sequence 95, Appl
30	315.5	12.7	1564	4	US-10-144-198-2	Sequence 2, Appl1
31	315.5	12.7	1564	4	US-10-144-198-4	Sequence 4, Appl1
32	315	12.7	647	5	US-10-691-672A-3	Sequence 3, Appl1
33	315	12.7	1960	4	US-10-236-031B-62	Sequence 62, Appl
34	315	12.7	1960	4	US-10-028-248A-104	Sequence 104, App
35	315	12.7	1960	4	US-10-107-782-104	Sequence 49, Appl
36	315	12.7	1960	5	US-10-696-909A-49	Sequence 27, Appl
37	315	12.7	1960	6	US-11-009-554-27	Sequence 32075, A
38	315	12.7	2143	5	US-10-450-763-32075	Sequence 36, Appl
39	313.5	12.6	1959	4	US-10-028-248A-36	Sequence 36, Appl
40	313.5	12.6	1959	4	US-10-107-782-36	Sequence 106, App
41	312.5	12.6	1959	4	US-10-028-248A-106	Sequence 106, App
42	312.5	12.6	1959	4	US-10-107-782-106	Sequence 106, App
43	312.5	12.6	1985	5	US-10-732-923-3351	Sequence 3351, Ap
44	311.5	12.5	2056	6	US-11-097-143-4824	Sequence 4824, Ap
45	311	12.5	240	4	US-10-101-487-75	Sequence 75, Appl

ALIGNMENTS

```
RESULT 1
US-10-732-923-3341
; Sequence 3341, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OR INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3341
; LENGTH: 1909
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-732-923-3341
Alignment Scores:
Pred. No.: 4.53e-156 Length: 1909
Score: 2406.50 Matches: 487
Percent Similarity: 89.89% Conservative: 2
Best Local Similarity: 89.52% Mismatches: 4
Query Match: 96.80% Indels: 51
DB: Query Match: 5 Gaps: 2
US-09-837-344-41 (1-1482) x US-10-732-923-3341 (1-1909)
Oy 1 CAAGAACACAAAGCCGATCTAGAACACAGACAGCGTCTAAAGAAAAGTTGCAGAACAA 60
Db GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1365
Oy 61 CAAGCATTTAGAACAGATAGACTT----- 87
Db GlnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 1405
```

QY 88 -----GCTAAAGAAAAGTTACAAAGCAGCAAGGATTAGAACAA 129
Db 1406 LeuGInGInuArGArGAlaLySGluLySLeuGInGInSeRAspLeuGIn 1425
QY 130 GAGACACTTGCTAAAGAAAAGTTGCAGAACAAACGATCTAGAACAAAGAGAGCT 189
Db 1426 GluArGArGAlaLySGluLySLeuGInGInSeRAspLeuGInGInuArGrG 1445
QY 190 GCTAAAGAAAAGTTGCAGAACAAACGATCTAGAACAAAGAGAGAGCTCTAAAGAA 249
Db 1446 AlaLySGluLySLeuGInGInSeRAspLeuGInGInuArGArGAlaLySGlu 1465
QY 250 AAGTTGCAGAACAAACGATCTAGAACAAAGTTCGCTTAAGAAAAGTTACAA 309
Db 1466 LySLeuGInGInuArGArGAlaLySGluLySLeuGInGInuArGrGAla 1485
QY 310 GAGCAGCAAGCGATTAGAACAAAGAGAGAGCTGCTAAAGAAAAGTTGCAGAACAA 369
Db 1486 GluGInGInuArGArGAlaLySGluLySLeuGInGInuArGrGAlaLySGlu 1505
QY 370 AGCGATTTAGAACAAAGAGAGAGCTGCTAAAGAAAAGTTGCAGAACAAACGATTTA 429
Db 1506 SerAspLeuGInGInuArGArGAlaLySGluLySLeuGInGInuArGrGAla 1525
QY 430 GACCAAGAGAGAGCTGCTTAAGAAAAGTTGCAGAACAAACGATTTAGAACAAAG 489
Db 1526 GluGInGInuArGArGAlaLySGluLySLeuGInGInuArGrGAlaLySGlu 1545
QY 490 AGACGTGCTTAAGAAAAGTTGCAGAACAAACGATTTAGAACAAAGAGAGAGCTGCT 549
Db 1546 ArGArGAlaLySGluLySLeuGInGInuArGArGAlaLySGluLySLeuGInGInuArGrGAla 1565
QY 550 AAAAGAAAAGTTGCAGAACAAACGATTTAGAACAAAGAGAGAGCT----- 597
Db 1566 LySLeuGInGInuArGArGAlaLySGluLySLeuGInGInuArGrGAlaLySGluLyS 1585
QY 597 ----- 597
Db 1586 LeuGInGInuArGArGAlaLySGluLySLeuGInGInuArGrGAlaLySGluLySLeuGInGInu 1605
QY 598 -----GCTAAAGAAAAGTTGCAGAACAGCAGCAAGAA 627
Db 1606 GluGInGInuArGArGAlaLySGluLySLeuGInGInuArGrGAlaLySGluLySLeuGInGInuAr 1625
QY 628 GATTAGAACAAAGAGAGCTGATACGAAAAAATTTAGAAAAAGAAAAAGAACATGGA 687
Db 1626 AspLeuGInGInuArGArGAlaLySGluLySLeuGInGInuArGrGAlaLySGluLySLeuGInGInu 1645
QY 688 GATATATTAGCAGAGATTTATATGTCGTTTAAAGAAATCCAGCTATTAACCTCCATCA 747
Db 1646 AspValLeuAlaGluAspLeuLyGlyArGrGAlaLySGluLySLeuGInGInuArGrGAlaLySGluLyS 1665
QY 748 GAAATAGAGCGATATATATACACATCAATCTTTTACCTCAGAGAACAGAGGG 807
Db 1666 GluAsnGluArGlyTyTylleProHISGInSeRSeuLeuProGlnAspAsnArGly 1685
QY 808 AATAGTAGAGATTCCAGAGAAATATCTATTAAGAAAAACAATAGAGAAATCTATTACA 867
Db 1686 AsnSerArGAspSerLySLeuIleSerIleIleGluLySThrAsnArGlyIleSerIleThr 1705
QY 868 ACAATGTTGAAAGAGAGAGATATATACATAAGAAACATCTTGAGAAAAAGAAAGTGT 927
Db 1706 ThrAsnValGluGlyArGrGAlaLySGluLySLeuGInGInuArGrGAlaLySGluLySLeuGInGInu 1725
QY 928 TCAATATAACAGAAACAAAGAAAGTAATCTGCAGATACAAATCATCATTAAGAG 987
Db 1726 SerIleLySProGInuLySGluLySLeuAspIleGlnAsnHISThrLeuGlu 1745
QY 988 ACAAGTAATATTTCTGATGTTAATGATTTTCAATAGATAGATGAGATGAATTAAGT 1047
Db 1746 ThrValAsnIleSerAspValAsnAspHeGlnIleSerLySLeuAspGluIleSer 1765

QY 1048 GCTGAATATGACGATTCATTAATAGATGAGAGAAAGATGATGAGACTTAGACGAATTT 1107
Db 1766 AlaGluTyRAspAspSerLeuIleAspGluGluAspAspGluAspLeuAspGluPhe 1785
QY 1108 AAGCCTATTTGCAATATGACAAATTTCCAGATGAGAGAAAACTATGAAATTTATTAAGAA 1167
Db 1786 LysProIleValGlnTyRAspAsnPheGlnAspGluAsnIleGlyIleTyRLeuGlu 1805
QY 1168 CTAGAGATTTGATAGAGAAAAATGAAAAATTTAGATGATTTAGATGAGGAATAGAAAAA 1227
Db 1806 LeuGluAspLeuIleGluLySAsnGluAsnLeuAspAspLeuAspGluGlyIleGluLyS 1825
QY 1228 TCATCAGAGAAATTAATCTGAAGAAAAATTAAGAAAAAGGAAATATGAAAAAACAAG 1287
Db 1826 SerSerGluGluLeuSerGluGluLySleLySLeuGlyLySLeuTyRLeuLySLeuTyR 1845
QY 1288 GATATAATTTTAAACCAATGATTAATAAGTTGTATGATGAGCATATTTAAAAAATATATA 1347
Db 1846 AspAsnAsnPheLySProAsnAspLySLeuTyRAspGluHISleLySLeuTyRLeuLyS 1865
QY 1348 AATGATTAAGCGATTAATAAGAAAAAGGAAAAATTCATTAATACTGTTTCATATATTT 1407
Db 1866 AsnAspLySGlnValAsnGluLySLeuGluLySLeuPheIleLySLeuPheHISlePhe 1885
QY 1408 GACGAGACAATGAATTTTACAGATCGTGATGAGATTTATCTGAAGATATACTAAATAT 1467
Db 1886 AspGlyAspAsnGluIleLeuGlnIleValAspGluLeuSerGluAspIleThrLySLeuTyR 1905
QY 1468 TTTATGAAACTA 1479
Db 1906 PheMetLySLeu 1909
RESULT 2
US-09-837-344-31
; Sequence 31, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:

Db 1 SerhapleuglIngluhArghalalyseglulyleugInglInglInserhpeu 20
Qy 73 GAACAAGATAGACTTGCTAAAGAAAAGTTACAGAGCGCAAGCGATTGTAGAACAGAG 132
Db 21 GlInghapArghleuAlalyseglulyleugInglInglInserhpeuInglInglu 40
Qy 133 AGACTTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGAGCGTGT 192
Db 41 ArghalalyseglulyleugInglInglInserhpeuInglInglIngluhArghal 60
Qy 193 AAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGTAAAGAAAAG 252
Db 61 LyseglulyleugInglInglInserhpeuInglIngluhArghalalyseglulys 80
Qy 253 TTGCAAGAACAAAGCGATTGTAGAACAGTCTGCTAAAGAAAAGTTACAGAG 312
Db 81 LeugInglInglInserhpeuInglInghArghleuAlalyseglulyleugInglu 100
Qy 313 CAGCAAGCGATTGTAGAACAGAGCGTGTAAAGAAAAGTTGCAAGAACAGAGC 372
Db 101 GlInghInserhpeuInglIngluhArghalalyseglulyleugInglInglInser 120
Qy 373 GATTAGAACAAAGAGAGCGTGTAAAGAAAAGTTGCAAGAACAGAGCGATTAGAA 432
Db 121 AsepleugInglIngluhArghalalyseglulyleugInglInglInserhpeuIn 140
Qy 433 CAAGAGAGCTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAG 492
Db 141 GlIngluhArghalalyseglulyleugInglInglInserhpeuInglIngluhAr 160
Qy 493 CGTGTAAAGAAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGCGTGTAA 552
Db 161 ArghalalyseglulyleugInglInglInserhpeuInglIngluhArghalalyse 180
Qy 553 GAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGTAAAGAAAAGTTG 612
Db 181 GlulyleugInglInglInserhpeuInglIngluhArghalalyseglulysleu 200
Qy 613 CAAGAGCGCAAGAGATTTAGAACAAAGAGCGTGATACGAAAAAAATTTAGAAAGA 672
Db 201 GlInglInglInghArghleuInglInghArghleuInglInghArghleuInglu 220
Qy 673 AAAAGGAGCATGAGATATTTAGAGAGAGATTTATGTCGTTAAATAACAGCT 732
Db 221 LyseglulyleuInglInglInserhpeuInglIngluhArghleuInglInglIn 240
Qy 733 ATAGACTTCATCAGAAAAATGAAAGCTGATATTATATACCATCAATCTTCTTA 792
Db 241 IlleglInleuProsergluhIngluhArghalalyseglulyleuInserleuPro 260
Qy 793 CAGGACACAGAGGGAATGTAGAGATTTCCAGGAAATTCCTTAATAAGAAAAACAA 852
Db 261 GlnAseArghleuInglInghArghleuInglInghArghleuInglInghArghleu 280
Qy 853 AAGAGATCTATTACAAAGATTTGAAAGCGAGGAGATATCATTAAGAGACATCTTGA 912
Db 281 ArghleuInglInghArghleuInglInghArghleuInglInghArghleuInglu 300
Qy 913 GAAAAGAAAGATGTTCAATAAACAGAACAAAGAAAGATTAATCT 960
Db 301 GlulyleuInglInghArghleuInglInghArghleuInglInghArghleuInglu 316

RESULT 4
US-09-837-344-20
; Sequence 20, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathie
STREET: P. O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-837-344-20
Alignment Scores:
Pred. No.: 4,35e-33 Length: 117
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.26% Indels: 0
Gaps: 0
US-09-837-344-41 (1-1482) x US-09-837-344-20 (1-117)
Qy 610 TTGCAAGAGCGCAAGAGATTTAGAACAAAGAGCGTGATACGAAAAAAATTTAGAA 669
Db 1 LeugInglInglInghArghleuInglInghArghleuInglInghArghleuInglu 20
Qy 670 AGAAAAGGAGCATGAGATATTTAGAGAGATTTATGTCGTTAAATAACCA 729
Db 21 ArghleuInglInghArghleuInglInghArghleuInglInghArghleuInglu 40
Qy 730 GCTATAGACTTCATCAGAAAAATGAAAGCTGATATTATATACCATCAATCTTCTT 789
Db 41 AlalleglInleuProsergluhIngluhArghalalyseglulyleuInserleuPro 60
Qy 790 CCTCAGACACAGAGGGAATGTAGAGATTTCCAGGAAATTCCTTAATAAGAAAAACA 849
Db 61 ProGlnAseArghleuInglInghArghleuInglInghArghleuInglInghArghleu 80
Qy 850 AATAGAGATCTATTACAAAGATTTGAAAGCGAGGAGATATCATTAAGAGACATCTT 909
Db 81 AanaArghleuInglInghArghleuInglInghArghleuInglInghArghleuInglu 100
Qy 910 GAAAGAAAAGATGTTCAATAAACAGAACAAAGAAAGATTAATCT 960

US-09-837-344-41 (1-1482) x US-09-837-344-19 (1-107)

QY 640 AGAAGGCTGATACGAAAAAATTAGAAAAAAGAAACATGAGATATATTAGCA 699
 |||||
 Db 1 ArglysaLaApThrlYsLysaSnLeuGluArgLyLysGluHlsGlyAspIleuAla 20
 QY 700 GAGGATTTATATGTCGTTTAGAAATACAGCATATGAACTTCCATCAGAAAAATGAACT 759
 |||||
 Db 21 GluApleuTyrgClYArgLeuGluIleProAlaIleGluLeuProSerGluuSnGluArg 40
 QY 760 GGATATTATATACCATCATCTTCTTTACCTCAGACACACAGAGGGAATGTAGAGAT 819
 |||||
 Db 41 GlyTyTyTyIleProHlsGlnSerSerLeuProGlnAspAsnArgGlyAsnSerArgAsp 60
 QY 820 TCCAGGAAATATCTATATAGAAAAAACAATAGATCTTTTACCAAAATGTTGAA 879
 |||||
 Db 61 SerLySgLuIleSerIleIleGluYsThrAsnArgGlnSerIleThrThrAsnValGlu 80
 QY 880 GAGCAGAGGATATACATTAAGACATCTTGAAGAAAGAAAGATGTTCAATTAACCA 939
 |||||
 Db 81 GlyArgArgAspIleHlsYsGlyHlsLeuGluGluYsLysAspGlySerIleYsPro 100
 QY 940 GAACAAAAAGAAATTAATCT 960
 |||||
 Db 101 GluGlnLySgLuAspLySer 107

RESULT 7
 US-09-900-963-19
 ; Sequence 19, Application US/09900963
 ; Publication No. US20030064075A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DRUILHE, Pierre
 ; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
 ; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
 ; OF STIMULATING THE T LYMPHOCYTES
 ; NUMBER OF SEQUENCES: 46
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis
 ; STREET: P.O. Box 1404
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: United States
 ; ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/900,963
 ; FILING DATE: 10-Jul-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/098,327
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McGowan, Malcolm K.
 ; REGISTRATION NUMBER: 39,300
 ; REFERENCE/DOCKET NUMBER: 010830-045
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 836-6620
 ; TELEFAX: (703) 836-2021
 ; INFORMATION FOR SEQ ID NO: 19:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 107 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; PUBLICATION INFORMATION:
 ; DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
 ; US-09-900-963-19

Alignment Scores:
 Pred. No.: 9,54e-30 Length: 107
 Score: 554.00 Matches: 107
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 22.28% Indels: 0
 DB: 3 Gaps: 0

US-09-837-344-41 (1-1482) x US-09-900-963-19 (1-107)

QY 640 AGAAGGCTGATACGAAAAAATTAGAAAAAAGAAACATGAGATATATTAGCA 699
 |||||
 Db 1 ArglysaLaApThrlYsLysaSnLeuGluArgLyLysGluHlsGlyAspIleuAla 20
 QY 700 GAGGATTTATATGTCGTTTAGAAATACAGCATATGAACTTCCATCAGAAAAATGAACT 759
 |||||
 Db 21 GluApleuTyrgClYArgLeuGluIleProAlaIleGluLeuProSerGluuSnGluArg 40
 QY 760 GGATATTATATACCATCATCTTCTTTACCTCAGACACACAGAGGGAATGTAGAGAT 819
 |||||
 Db 41 GlyTyTyTyIleProHlsGlnSerSerLeuProGlnAspAsnArgGlyAsnSerArgAsp 60
 QY 820 TCCAGGAAATATCTATATAGAAAAAACAATAGATCTTTTACCAAAATGTTGAA 879
 |||||
 Db 61 SerLySgLuIleSerIleIleGluYsThrAsnArgGlnSerIleThrThrAsnValGlu 80
 QY 880 GAGCAGAGGATATACATTAAGACATCTTGAAGAAAGAAAGATGTTCAATTAACCA 939
 |||||
 Db 81 GlyArgArgAspIleHlsYsGlyHlsLeuGluGluYsLysAspGlySerIleYsPro 100
 QY 940 GAACAAAAAGAAATTAATCT 960
 |||||
 Db 101 GluGlnLySgLuAspLySer 107

RESULT 8
 US-10-732-923-3319
 ; Sequence 3319, Application US/10732923
 ; Publication No. US20050108791A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Edgerton, Michael D
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
 ; FILE REFERENCE: 38-15(52796)C
 ; CURRENT APPLICATION NUMBER: US/10/732,923
 ; PRIOR FILING DATE: 2003-12-10
 ; PRIOR APPLICATION NUMBER: 10/310,154
 ; NUMBER OF SEQ ID NOS: 24149
 ; SEQ ID NO 3319
 ; LENGTH: 1003
 ; TYPE: PRT
 ; ORGANISM: Kaposis's sarcoma-associated herpesvirus
 ; US-10-732-923-3319

Alignment Scores:
 Pred. No.: 5,35e-18 Length: 1003
 Score: 381.50 Matches: 101
 Percent Similarity: 53.41% Conservative: 126
 Best Local Similarity: 23.76% Mismatches: 105
 Query Match: 15.35% Indels: 93
 DB: 5 Gaps: 15

US-09-837-344-41 (1-1482) x US-10-732-923-3319 (1-1003)

QY 4 GAACAAACAAGCGATCTAGAACAGAGACGTCGTTAAGAAAGTTGCAAGAACAA 63
 |||||
 Db 440 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 459
 QY 64 AGCAGTTTGAACAAGATGACTTCTTAAGAAAAAGTTTCAAGAGCAGCAAGCGATT 123
 |||||

Db	460	gIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	479
Qy	124	GAACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAGAACAAAGCGATCTAGAACAAAG	183
Db	480	gIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	499
Qy	184	AGAGCTGCTTAAGAAAAAGTTGCAGAGAACAAAGCGATTTAGAACAAAGAGAGCTGCT	243
Db	500	gIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	518
Qy	244	AAAGAAAAGTTGCAGAGAACAAAGCGAT-----TTAGAACAAAGATAGACTTGTCTAA	297
Db	519	gIngnIngnIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIngnIuPPrognIngnIngnIuPPrognIn	538
Qy	298	GAAGAAATTACAGAGAGCAGCAAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTG	357
Db	539	gIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	558
Qy	358	CAAGAACCAAAAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAGAGAACAA	417
Db	559	gIngnIngnIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIngnIuPPrognIngnIngnIuPPrognIn	575
Qy	418	CAAGCGATTTAGAACAGAGAGAGCTTCTAAAGAAAAGTTGCAGAGAACAAAGCGAT	477
Db	576	gIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	592
Qy	478	TTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAGAGAACAAAGCGATTTAGAACAA	537
Db	593	LeuIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	609
Qy	538	GAGAGAGCGTCTTAAGAAAAGTTGCAGAGAACAAAGCGATTTAGAACAGAGAGCGT	597
Db	610	gIn-----gIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	624
Qy	598	GCTAAAGAAAAGTTGCAGAGAGAACAAAGATTTAGAACAAAGAGAGCGTATAGCAAA	657
Db	625	--gIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	643
Qy	658	AAAAATTTAGAAAAGAAAAGAACATGAGATATATTAGCAGAGAGATTTATATGTCGT	717
Db	644	gIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	662
Qy	718	TTAGAAAATACAGACTATAGAACTTCCATCAGAAAAATGAAAGTATATTATATACACAT	777
Db	663	LeuIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	676
Qy	778	CAATCTTTTAACCTCAGGACACAGAGAGGAATAGTAGAGATTCCAGAGAAATATCTATA	837
Db	676	-----	676
Qy	838	ATTAGAAAACAAATAGAGATCTATTACAACAAATGTTGAAGAGAGATATACAT	897
Db	677	LeuIngnIngnIngnIngnIu-----	683
Qy	898	AAAGAGACATCTTGAAGAAAAGAACATGTTCAATTAACAGAACAAAGAAAGATATAA	957
Db	684	-----LeuIngnIngnIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	699
Qy	958	TCTGCTGACATACAAATTCATACATTAGAGACAGTAAATATTCTTCGATGTTATGATTTT	1017
Db	700	-----LeuIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	706
Qy	1018	CAAAATTAAGTATGATGAGATGAATTAAGTGTGAATTAAGCGATTAATATAGATGA	1077
Db	707	gIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	726
Qy	1078	GAGAGAGATGATGAGACTTAGACCAATTAAGCCTATGTGCAATATAGCAATTTCCAA	1137
Db	727	gIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	738
Qy	1138	GATGAGAGAAACATAGAAATTTATATAAGACTTAAGAAATTTGATAGAGAAAATGAA---	1199
Db	739	gIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIngnIngnIuPPrognIn	753

```

QY      1195 AATTAGATGATTTAGATGAA-----CGATTAGAAAAATCATCGAAGATT 1242
Db      754 Gltuengtugluvalgtuglucnglucnglncglvalgtuglucnglucntr 773
QY      1243 TCTGAGAAAAAATA 1257
Db      774 ValgtugluProlle 778

RESULT 9
US-10-732-923-3359
; Sequence 3359, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3359
; LENGTH: 1407
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-732-923-3359

Alignment Scores:
Pred. No.:      6,23e-18      Length:      1407
Score:          380.50      Matches:      127
Percent Similarity: 45.84%      Conservative: 132
Best Local Similarity: 22.48%      Mismatches:  167
Query Match:    15.31%      Indels:      139
DB:              5      Gaps:      20

US-09-837-344-41 (1-1482) x US-10-732-923-3359 (1-1407)
QY      1 CAAAGAACCAAAAGCGATCTA-----GAACAAGAGAGACGTCCTAAAGAAAGTTG 51
Db      140 GtngtuglucngluarvgltuLeuAlaGltuglucnglucnglncArgrvLbVArvgluarPhe 159
QY      52 CAAGAA-----CAACAAAGCGATTTAGAACAAAGATGACTTGTCTAAAGAAAG 99
Db      160 GltuglnhstYrserArxgrIntYArxAspLysglucn---ArxleuGlnarGlnglu 178
QY      100 TTACAAGAGCAGCAAAAGCGATTTAGAACAAAGAGACTTGTCTAAAGAAAGTTGCCAAGAA 159
Db      179 LeugltugluarvgArAla-----GltuglucnglucnLeuArvgArvgLyGlyArghp 196
QY      160 CAACAAAGCGATCTAAGAACAGAG-----AGACGTGCTAAAGAAAGATTGCCAAGAA 210
Db      197 AlaGltugluPheIleGltuglucnglucnLeuArvgArvgArvgltuglucnglucnLeuYArgr 216
QY      211 CAACAAAGCGATTTAGAACAAAGAGAGACGTCCTAAAGAAAG-----TTG 255
Db      217 GltuLeuArvgltuglucnglucnGlnArvgArvgltuArvgArvgltuglnhAgluArGAlaLeu 236
QY      256 CAAGAACAAAGAGCGATTTAGAACAAAGATGACTTGTCTAAAGAAAGTTACAAAGACAG 315
Db      237 GtngtuglucnglucnglucnLeuLeuArvgltuArvgArvgArvgltugluPArArvgltugln 256
QY      316 CAAGCGATTTAGAACAAAGAGAGACGTCCTAAAGAAAGTTGCCAAGAACAAAGCAT 375
Db      257 Gln-----GlnLeuArvgArvgltuLeuGltuglnuLeuArvgltuArvgltuglnArgr 272
QY      376 TTAGAACAAAGAGAGACGTCCTAAAGAAAGTTGCCAAGAACAAAGCGATTTAGAACAA 435
Db      273 LeugltuglucnglucnglucnArvgArvgltuglucnLeuArvgArvgltuglnArgr--LeuGltugln 291
QY      436 GAG-----AGACTTCTAAAGAAAGTTGCCAAGAACAAAGC 474
Db      292 GtngluArvgArvgltuglucnLeuArvgArvgltuLeuGltuglnuLeuArvgltuLeuArvgltugln 311

```

[illegible][illegible]

Oy	394	-----GCTAAAGAAAAGTTCGACAAACAAACAGCGATTATGAAACAAG	438
Db	377	ArgTyrThrLeuTyralaIysProalaglmarGlnGlnValargGlnGlnGlnLeu	396
Oy	439	AGACTTGCTAAAGAAAAGTTCGACAAACAAACAGCGATTATGAAACAAGAGA	492
Db	397	ArgLeuIysbGlnGlnIuLybLeuGlnIhargGlnIuLybArgargGlnIuLybGlnIuArgGlnTyr	416
Oy	493	-----CGTGTAAAGAAAAGTTCGACAAACAAACAGCGATTATGAA	534
Db	417	ArgGlnIuValGlnLeuGlnIhargGlnIuArgGlnIuArgGlnIhargGlnIuGlnLeuGln	436
Oy	535	CAAGAGACGCTGCTAAAGAAAAGTTCGACAA	567
Db	437	ArgGlnIuValargGlnIuLybArgargGlnIuGlnIuArgGlnIuLybGlnIuVal	456
Oy	568	-----CAACAAGACGATTATGAAACAAAGACACGCTGTAAAGAAAAGTTCGACAG	618
Db	457	GlnLeuTrrGlnIuArgGlnIuGlnIhLeuGlnIhargGlnIuArgGlnIuLybArgargGlnIuArg	476
Oy	619	CAGCAAGACAGATTATGAA	648
Db	477	GlnIuYsbGlnTyrLeuGlnIuLybValGlnLeuIhargGlnIuGlnIuGlnLeuGlnIhargGlnIu	496
Oy	649	GATACGAAAAAAATTTAGAAAAGAAAAAG-----GAACATCGAGATATATTTAGCA	699
Db	497	ArgGlnIuYbArgargGlnIuArgGlnIuArgGlnTyrLeuGlnIuLybValGlnLeuGlnIu	516
Oy	700	GAGGATTTATATGGTCGTTTGAATAACACACTATAGAACTTCATCGATCGAAAAATGAAGCT	759
Db	517	GlnGlnGlnIuLeuGlnIhargGlnIu-----ArgGlnIuYbArgargGlnIuArgGlnIuArg	534
Oy	760	GGATATTAT-----ATACACATCATCTTTTACTCTAGAGCAACAGAGGG	807
Db	535	GlnTyrLeuGlnIuLybValGlnLeuGlnIuGlnIuGlnIuGlnIuGlnIhargGlnIuArgGlnIu	554
Oy	808	AATAGTACAGATTTCCACAGAAAATATCTATATAGAAAAAACAAATAGAGAA-----	858
Db	555	LybArgargGlnIuArgGlnIuYsbGlnTyrLeuGlnIuLybValGlnLeuGlnIuGlnIu	574
Oy	859	-----TCTATTACAAACAATGTTGAAGACAGGAAGGGATATACATTAAGACATCTTGAA	912
Db	575	GlnLeuGlnIhargGlnIuArgGlnIuLybArgargGlnIuArgGlnIuLybGlnTyrLeuGlnIu	594
Oy	913	GAA-----AAGAAAGATCGTTCAATTAATAAACCAAGCAAAAAAGAGATTAATTCGCT	963
Db	595	LybValGlnLeuGlnIuGlnIuGlnIuGlnIuGlnIhargGlnIuArgGlnIuLybArgargGlnIu	614
Oy	964	GACATACAAAATCATACATTAAGACAGCACTAAATATTTCTGATGTTAATGATTTTCAA--	1020
Db	615	GlnArgargIuArgGlnTyrLeuGlnIuLybValGlnLeuGlnIuGlnIuGlnIuValGlnIharg	634
Oy	1021	-----ATAAGT	1026
Db	635	GlnGlnIuArgGlnIuLybArgargGlnIuArgGlnIuArgGlnTyrLeuGlnIuLybGlnLeuGlnIu	654
Oy	1027	AAGTATGAGATGAATTAAGTCGTGATGATATGACGATTCAATTAATGATGAGAAAGAAAT	1086
Db	655	ArgGlnIuGlnIuArgLeuGlnIuGlnIu-----GlnGlnIuLeuLeuArgGlnIuArgGlnIu	672
Oy	1087	GATGAAGACCTTAGACGAATTTAAGCCTTATGTGCAATATGACAAATTTCCAAAGATGAAGAA	1146
Db	673	LybArgargGlnIuArgGlnIuArgGlnIuTyrLeuGlnIuLybValGlnLeuGlnIuGlnIuGlnIu	692
Oy	1147	AACATA-----GGAATTTATTAAGAA	1167
Db	693	GlnLeuGlnIhargGlnIuArgGlnIuLybArgargGlnIuArgGlnIuArgGlnIuTyrLeuGlnIu	712
Oy	1168	CTAGAAAGATTTGATAGACAAAATGAAAAATTTAGATGATGATTAAGATGAGCAATGAAAAA	1227
Db	713	LybGlnIuGlnIuLeuGlnIhargGlnIuGlnIuArgLeuGlnIhargGlnIuLybGlnIuGlnIuGlnIharg	732
Oy	1228	TCATCAGAAAGAA-----TTATCTGAAAGAAAAATA	1257

```

Db      733  GlnbapargGlnlybArgArgGlnValArgGlnlybTyLeuGlnlybGlnlybGln 752
Oy      1258  AAAAAAGGAAAGAAATATGAAAAAACAAGGATAATTTTAAACCAATGATTAAGT 1317
Db      753  GlnGlnGlnbapargbArgbGlnbArgGlnlybGlnleuLeuArgGlnbapargbArg 772
Oy      1318  TTGTTATGATGAGCATTTTAAAAAATATATAAATGATTAAGCAGTTTAAATGAAGAA 1377
Db      773  GlnTyLeuGlnlybValGlnleuGlnbArgGlnGlnGlnGlnGlnGlnlybArg 792
Oy      1378  AAA 1380
Db      793  Arg 793

RESULT 11
US-10-450-763-47660
; Sequence 47660, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO: 47660
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (362)..(408)
; OTHER INFORMATION: Vertebrate metallochioneine proteins domain identified by
; OTHER INFORMATION: EMATRIX, accession number BL00203, p-value=2.484e-11, raw score
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(419)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-47660

Alignment Scores:
Pred. No.: 1,19e-17 Length: 419
Score: 376.50 Matches: 149
Percent Similarity: 44.34% Conservative: 35
Best Local Similarity: 35.90% Mismatches: 76
Query Match: 14.26% Indels: 155
DB: 5 Gaps: 25

US-09-837-344-41 (1-1482) x US-10-450-763-47660 (1-419)
Oy      1217  CCTGATCATTAATCATCTAAATTTTCTATCATCAAAATCTTACTTATTA 1158
Db      136  PTOALPPTOALSerSerProCysSer---ArgSerLeuSerSerSerArgSerLeu--- 153
Oy      1157  ATTCTCATGTTTCTTCTCATCTTGAAA-----TTGTCATATTGCACA 1116
Db      154  -----SerSerSerCysArgValSerCysGlnHiserLeuSerSerCysCys 169
Oy      1115  ATAGCGTTAAATTCGCTGAAGCTTTCATCATCTTCTTCTTCATCTTAATGAATGCTCA 1056
Db      170  ArgSerLeuSerCysSerArgSerSerSerThrSerSerArgSerSerSerThrSerSer 189
Oy      1055  TATTGAGCAGCTAATTATTCATCCCTACTTACTTATTGAAAAATCAATTAACATCAGAAATA 996

```

```

Db      190 SerSerLeuSerLeuSerSerLeu-----198
Oy      995 TTAACTGCTCTAAATGATGATTGGATGACAGAAATTATCTTTGTTGCTGT 936
Db      199 -----SerCysSerSerCysArgLeuSer-----CysSer---208
Oy      935 TTATATGAAACATCTTTCTTTTCTTCAAAATGCTCTTATGATATGCCCTTCGCTTCA 876
Db      208 -----208
Oy      875 ACAATTGGTATAGATATCTCTATTGTTTTCTATATAGATATTCCTTGGAATCT 816
Db      208 -----208
Oy      815 CTACTATTCCTCTGTGTTCTCGAGGTAAAGAAATGATGTGTATATTAATTCACGT 766
Db      209 -----Asn 209
Oy      755 TCATTTTCTGATGGAAGTTCTAATAGCTGATTTCTTAAACACATATTAATTCCTGTCT 696
Db      210 SerPheSerThrSerSerSerLeuSerLeuSer-----221
Oy      695 AATATATCTCAATGTCCTTTTCTTTCTTAAATTTTTTGGATACAGCTTCCTTTGT 636
Db      222 -----SerSerLeuSerCys 226
Oy      635 TCTAAATCTCTTGTGCTGCTCTTGGCAACTTTTCTTAAACAGCTCTCTGTGTTCTAAATCG 576
Db      227 Ser-----SerCysCysLeuSerCysSerArganSerPheSerThrSerSerSerLeuSer 245
Oy      575 CTGTGTGTTCTTGGCAACTTTTCTTAAACAGCTCTCTGTGTTCTTAAATCGCTTTGTTGT 516
Db      246 LeuSerSerSerLeuSerCysSerSerCysArgLeuSerCysSerArganSerPhe-----263
Oy      515 TCTTGAACATTTTCTTAAACAGCTCTCTGTGTTCTTAAATCGCTTTGTTCTTGAAC 456
Db      264 -----SerThrSerSerSerSerLeuSerCysProArgSerPheSerCysSerHisSer 281
Oy      455 TTTTCTTTAGCAAGTCTCTGTTGTTCTTAAATCGCTTTGTTCTTGAACATTTTCTTAA 396
Db      282 LeuSer-----SerCysSerArgSerLeu-----SerSerCysSerHisSerLeu 296
Oy      395 GCACGTCTCTGTTGTTCTTAAATCGCTTTGTTGTTCTTGAACATTTTCTTAAACAGCTC 336
Db      297 Ser-----SerCysSerArgSerLeu-----SerSerCysSerHisSerLeuSer-----311
Oy      335 TCTGTGTTAAATCGCTTCTGCTCTGTTGAATCTTTTCTTAAACAGCTATCTGTTCT 276
Db      312 SerCysSerArgSerLeu-----SerSerCysSerHisSerLeuSer-----SerCysSer 326
Oy      275 AAATCGCTTGTGTTGTTCTTGAACATTTTCTTAAACAGCTCTCTGTGTTCTTAAATCGCT 216
Db      329 ArgSerLeu-----SerSerCysSerHisSerLeuSer-----SerCysSerArgSerLeu 345
Oy      215 ---TGTTGTTCTTGAACATTTTCTTAAACAGCTCTCTGTGTTCTTAAATCGCTTTGTT 156
Db      346 SerSerCysSerCysSerLeuSer-----SerCysSerArgSerLeu-----Ser 360
Oy      158 TCTTGAACATTTTCTTAAACAGCTCTCTGTGTTCTTAAATCGCTTGTGCTGCTTGAAC 99
Db      361 SerCysSerHisSerLeuSer-----SerCysSerHisSerLeu-----SerSerCysSer 377
Oy      98 TTTTCTTTAGCAAGCTATCTATCTGTTCTAAATCGCTTTGT-----TGT 57
Db      378 ArgSerLeuSerSerCysSerCysSerLeuSerSerCysSerArgCysLeuSerSerCys 396
Oy      56 TCTTGAACATTTTCTTAAACAGCTCTCTGTGTTCTTAAATCGCTT 12
Db      398 SerCysSerLeuSer-----SerCysSerArgSerLeu 408

RESULT 12
US-10-754-079-157
; Sequence 157, Application US/10754079

```

```

/ Publication No. US2005000602A1
/ GENERAL INFORMATION:
/ APPLICANT: Hirai, Yoheti
/ TITLE OF INVENTION: OLIGOPEPTIDES FOR PROMOTING HAIR GROWTH
/ FILE REFERENCE: 46739200022
/ CURRENT APPLICATION NUMBER: US/10/754, 079
/ CURRENT FILING DATE: 2004-01-07
/ PRIOR APPLICATION NUMBER: US 10/155,922
/ PRIOR FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: PCT/Jp01/04691
/ PRIOR FILING DATE: 2001-06-04
/ PRIOR APPLICATION NUMBER: JP 2000-166903
/ PRIOR FILING DATE: 2000-06-05
/ PRIOR APPLICATION NUMBER: JP 2001-347340
/ PRIOR FILING DATE: 2001-11-13
/ PRIOR APPLICATION NUMBER: JP 2001-347338
/ PRIOR FILING DATE: 2001-11-13
/ PRIOR APPLICATION NUMBER: JP 2001-371175
/ PRIOR FILING DATE: 2001-12-05
/ PRIOR APPLICATION NUMBER: JP 2001-371136
/ PRIOR FILING DATE: 2001-12-05
/ PRIOR APPLICATION NUMBER: JP 2003-18991
/ PRIOR FILING DATE: 2003-01-08
/ NUMBER OF SEQ ID NOS: 162
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 157
/ LENGTH: 1439
/ TYPE: PRT
/ ORGANISM: Mus Musculus
US-10-754-079-157

Alignment Scores:
Pred. No.: 1.73e-17 Length: 1439
Score: 374.00 Matches: 130
Percent Similarity: 48.15% Conservative: 130
Best Local Similarity: 25.22% Mismatches: 170
Query Match: 15.04% Indels: 124
DB: 5 Gaps: 25

US-09-837-344-41 (1-1482) x US-10-754-079-157 (1-1439)
QY 1 CAAGAACAACAAAGCGATCTGAACAAGAGAGACGTCCTAAAGAAAAG----- 48
Db 597 GUGUGUUGLUYARXGLULEUDUGLNGLUAYGARVARGSLNGLINRGRARGLN 616
QY 49 ---TGCAGAAACAACAAGCGATTGTAAACAAGATAGACTTGCTTAA-----GAA 96
Db 617 ILEULSGUGUGLUGLNLPHGLINRGLUNHISGLINRGLUALARGRARGARGLN 636
QY 97 AAGTTACAAGACCAAA----- 114
Db 637 ThRPheGLUGLUGLUGLNLNGLUGLUGLSErARGrARGSLNGLINrGLU 656
QY 115 ---AACGATTTTGAACAAGAGACTTGCTTAAGAAAAGTTGCAADAACAAGCGAT 171
Db 657 GLYLVrPHeLeUGLUGLUGLNUrGLInLeUARGrThRGlnArGLUvGLInArGrArG 676
QY 172 CTAGAACAAGAGAGACGTGCTAAAGAAAAAG-----TTGCAADAACAAGCGAT 222
Db 677 GInGLUGLInGLUArGrLUrPhELInGLUGLUGLUnHISLeUDInGLInrGrLUrVGLU 696
QY 223 TTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCCAAGAACSAACAAAGCGATTTAGAACA 282
Db 697 LeUARGrGLInGLUCyARPARGrGLVSeRARGrGLInGLInrArgarGLInGLInArGrLU 716
QY 283 GATAGACTTGCTTAAAGAAAAGTTACAAGACGACGAAGCGATTTTAGAACAAGAGAGA--- 339
Db 717 GUGInLeUARGrArGrGLInGLU---ArGARGrInArGrPhEARGrArGrLUInGLInrGHIS 755
QY 340 -----CGTCCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGA--- 390
Db 736 LeUGlUvGrGLUGLUGLInGLUnLeUrArGrPARGrProSeLARGrArGrLUInGLInrGHIS 755

```

Oy	391	-----CCTGCTAAAGAAAGTTCGACAGAACAAACAAAGCCGATTATAGAACAAAGAGAGA----	441
Db	756	GIngluArxgIngluIngluInleuArxArpArProSerArxArxgIngluInArxHi	775
Oy	442	-----CTTGCTAAAGAAAGTTCGACAGAACAAACAAAGCCGATTATAGAACAAAGAGAGA----	492
Db	776	GIngluArxgIngluIngluInleuArxArpArProSerArxArxgIngluInArxHi	795
Oy	493	-----CCTGCTAAAGAAAGTTCGACAGAACAAACAAAGCCGATTATAGAACAAAGAGAGA----	543
Db	796	GIngluArxgIngluIngluInleuArxArpArProSerArxArxgIngluInArxg	815
Oy	544	-----CCTGCTAAAGAAAGTTCGACAGAACAAACAAAGCCGATTATAGAACAAAGAGAGAGCT	597
Db	816	LeuGlulArxgIngluIngluInleuArxArpArProSerArxArxgIngluInArxHi	835
Oy	598	GCTAAAGAAAGTTCGACAGACAGAACAAAGAGATTATAGAACAAAGAAAGCCGATTATAGACAA	657
Db	836	--GIngluArxgIngluIngluInleuArxArpArProSerArxArxg--GIngluIn	853
Oy	658	AAAAATTAGAAAGAAAAAGAAACATGAGATATATTAGACAGAGATTTATAGATGCTGCT	717
Db	854	ArxArxgInleuArxgIngluIngluIn-----LeuArxArpArxSerPheArxArxg	870
Oy	718	TTAGAAATACACAGCTATAGAACTTCATCAGAAAAATGAACGTGATATTATATACACAT	777
Db	871	GlulIn-----GlulEuArxArpArxgPheHiSgluIngluInArxg	886
Oy	778	CAATCTTTTATACCTCAGAGACAAACAGAGGAATAGTACAGT-----	819
Db	887	ArxgIngluInleuArxgIngluIngluInArxgIngluInArxArpArxgLeuArxValGluIn	906
Oy	819	-----	819
Db	907	GInleuArxgIngluInArxgIngluIngluInArxArxgIngluInCyArpArxgLeu	926
Oy	820	TCCAAAGAAATATCTATATAGAAAAAAACAATAGAGAA-----TTATATT	864
Db	927	HiArxgInleuGlnValArxgIngluInleuGlnuInArxgLeuArxArpArxgLeu	946
Oy	865	ACAACAAATGTTGAAGACGAAAGGATATACHTAAAGACATCTTGAAGAAAGAAAGAT	924
Db	947	ArxArxgIngluInleuArxArxArp--ArxLysPheHiSgluIngluInArxg--	964
Oy	925	GCTCAATAAAAACGAGCAAAAGAAAGTAATCTGCTGACATCAAAATCAT-----	978
Db	965	-----ArgHiSgluInleuPheHiSgluInleuArxgLeuGlnGlnuInProSerArxg	981
Oy	979	-----ACATTAGACAGATTAATTTCTGATGTTAATGAT	1014
Db	982	ArxPheArxArxgIngluIngluInleuArxgIngluInCyValGIngluInArxgLeuArxArp	1001
Oy	1015	TTTCAATTAAGTAACTATGAGATCAAAATAAGTCTGAATATGACGAT-----	1062
Db	1002	SerLysVilArxArxg--GIngluInleuArxArxgInuArxgIngluInArxgLeuArxg	1020
Oy	1063	-----TCATTAATTAAGTAAGAAAGAAAT	1086
Db	1021	AspArxgLeuVilArxArxArpGlnuInleuArxgIngluInleuArxgIngluIngluInleuArxg	1040
Oy	1087	GATGAGACCTTAGAC--GAATTTAAGCCTATTGTGCATATAGCAAAATTTCCAGATGAA	1143
Db	1041	ArxgIngluInleuArxArxgLeuPheArxgIngluIngluInleuArpGlnuInleuIngluIn	1066
Oy	1144	GAAGAACATA-----GGAATTTATAGAAAGCTAGAGATTTGATAGAC-----	1185
Db	1061	GluArxgLeuArxArpArxgLeuVilArxArxgIngluIngluInleuArxArxgIngluInleu	1080
Oy	1186	AAAAATGAAAAATTTAGATGATTTTAATAGAAAGAAATACATCAAGAAATATCT	1245
Db	1081	ArxArxgIngluIngluInleuArxArxgIngluIngluInleuArxArxgIngluInleuPheArxg	1100

```

QY      1246 GAAGAAAAAATAAATAAGAAAGAAAGAAATATGAAAAAACAAGAGTATATATTTAAACCA 1305
DB      1101 G|u|n|g|u|g|u|e|u|a|r|g|u|n|g|u|a|r|g|u|n|g|u|e|u|a|r|g|a|s|l|e|a|r|g| 1120
QY      1306 AATGAT-----AAAGTTGTATGATGAGCATTTTAAAAATATTAATAATGATPAAG 1356
DB      1121 Aa|g|u|n|t|u|e|u|a|r|g|u|n|g|y|e|u|e|g|u|n|g|u|n|t|u|e|u|a|r|g|u|n|g|u|a|r|g| 1140
QY      1357 CAGGTTAATAGAGAAAAAGAA 1377
DB      1141 Lysphear|g|u|n|g|u|n|g|u 1147

RESULT 13
US-10-101-487-58
; Sequence 58, Application US/10101487
; Publication No. US20020169125A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, DAVID W.
; APPLICANT: BERGMAN, PHILIP A.
; APPLICANT: LOFOUIST, ALAN
; APPLICANT: PIETZ, GREGORY E.
; APPLICANT: TOMPKINS, CHRISTOPHER K.
; APPLICANT: WAGGNER JR., DAVID W.
; TITLE OF INVENTION: RECOMBINANT PRODUCTION OF POLYANIONIC POLYMERS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 077319/0329
; CURRENT APPLICATION NUMBER: US/10/101,487
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/277,705
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide
US-10-101-487-58

Alignment Scores:
Pred. No.:      1,59e-16      Length:      350
Score:          360.00      Matches:      78
Percent Similarity: 50.24%      Conservative: 132
Best Local Similarity: 18.66%      Mismatches: 140
Query Match:    14.48%      Indels:      68
DB:             4      Gaps:      2

US-09-837-344-41 (1-1482) x US-10-101-487-58 (1-350)
QY      1 CAAGAGACAACAAGGATCTAGAAACAAGAGAGAGCGTGCTAAAGAAAGATTGTCAGAAACA 60
DB      1 G|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n| 20
QY      61 CAAGCGATTTTAGAAACAAGATAGACTTGCTTAAAGAAAGTTTCAAGAGACAGCAAGCGAT 120
DB      21 G|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n| 40
QY      121 TTAGAACAGAGAGACTTGCTTAAAGAAAGTTGCAAGACACAAAGCGATCTGAACAA 180
DB      41 G|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n| 60
QY      181 GAGAGACGCTGAAGAAAGAAAGTTGCAAGAGACAACAAGGCGATTTAGAACAGAGAGCGT 240
DB      61 G|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n| 80
QY      241 GCTAAAGAAAGTTGCAAGACAACAAGCGATTTAGAACAGATAGACTTGCTTAAAGAA 300
DB      81 G|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n| 100
QY      301 AAGTTACAGAGACGCAAGCGATTTAGACAACAAGAGAGAGCGTCTTAAAGAAAGTTGCA 360
DB      101 G|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n|g|u|n| 120

```

[illegible][illegible]

[illegible]

Search completed: December 7, 2005, 07:15:55
Job time : 154 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 7, 2005, 06:57:12 ; Search time 7 Seconds
(Without alignments)
2364.606 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 2486 1 CAAAGAACAAACGATCT.....AATATTATTGAACATATA 1482

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32527 seqs, 5584426 residues

Total number of hits satisfying chosen parameters: 65054

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xih
-O/cgn2_1/usfro.epool.b/us09837344/runat.05122005.133410.29585/app.query.fasta_1.1671
-DB=Published Applications_AA_New -OPMT=fastan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62
-TRANS=human40.cdt -LIST=45 -DCCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09837344 @CGN 1.1 @runat.05122005.133410.29585
-NCPU=6 -ICPU=3 -NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THRAD=5 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELop=6 -DELext=7

Database : Published Applications_AA_New:*

1: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pcp:*
2: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pcp:*
5: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pcp:*
7: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pcp:*
8: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pcp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	280.5	11.3	1404	US-10-878-556A-169	Sequence 169, App
2	266	10.7	793	US-11-060-914-2	Sequence 2, Appl
3	262.5	10.6	795	US-10-770-726-49	Sequence 49, Appl
4	254	10.2	1388	US-10-821-234-1143	Sequence 1143, Ap
5	251.5	10.1	718	US-10-878-556A-97	Sequence 97, Appl
6	241.5	9.7	885	US-10-793-626-1660	Sequence 1660, Ap
7	230.5	9.3	1189	US-11-074-176-134	Sequence 134, App
8	228	9.2	1586	US-10-821-234-901	Sequence 901, App
9	223	9.0	1410	US-10-878-556A-136	Sequence 136, App
10	220.5	8.9	989	US-10-821-234-975	Sequence 975, App

11	215.5	8.7	1410	US-10-821-234-1050	Sequence 1050, Ap
12	210.5	8.5	558	US-10-467-657-694	Sequence 694, App
13	210	8.4	1289	US-10-821-234-1145	Sequence 1145, Ap
14	208.5	8.4	898	US-11-099-691-7	Sequence 7, Appl
15	205	7.8	3500	US-11-085-775-2	Sequence 2, Appl
16	199.5	8.0	691	US-10-131-826A-16	Sequence 16, Appl
17	199.5	8.0	703	US-10-821-234-963	Sequence 963, App
18	195	7.8	5024	US-10-793-626-2964	Sequence 2964, Ap
19	190.5	7.7	664	US-10-878-556A-44	Sequence 44, Appl
20	190	7.6	531	US-11-060-914-4	Sequence 4, Appl
21	188	7.6	1095	US-10-793-626-3154	Sequence 3154, Ap
22	183	7.4	155	US-10-952-535A-17	Sequence 17, Appl
23	182	7.3	672	US-11-000-463-455	Sequence 455, App
24	181.5	7.3	1279	US-10-793-626-3188	Sequence 3188, Ap
25	181	7.3	1618	US-10-984-645-2	Sequence 2, Appl
26	180	7.2	145	US-10-952-535A-19	Sequence 19, Appl
27	180	7.2	1122	US-10-467-657-6112	Sequence 6112, Ap
28	179	7.2	284	US-10-821-234-1632	Sequence 1632, Ap
29	179	7.2	465	US-10-878-556A-116	Sequence 116, App
30	177.5	7.1	1616	US-10-821-234-1497	Sequence 1497, App
31	176.5	7.1	1565	US-10-467-657-2704	Sequence 2704, Ap
32	176	7.1	121	US-10-952-535A-14	Sequence 14, Appl
33	175.5	7.1	248	US-10-878-556A-175	Sequence 175, App
34	175	7.0	677	US-10-982-545-12	Sequence 12, Appl
35	174.5	7.0	747	US-10-821-234-1662	Sequence 1662, Ap
36	173.5	7.0	1596	US-11-060-005-4	Sequence 4, Appl
37	172.5	6.9	676	US-11-135-855-28	Sequence 28, Appl
38	172.5	6.9	717	US-11-135-855-29	Sequence 29, Appl
39	172.5	6.9	745	US-10-793-626-1500	Sequence 1500, Ap
40	172	6.9	1043	US-10-821-234-1055	Sequence 1055, Ap
41	167	6.7	468	US-10-131-826A-90	Sequence 90, Appl
42	167	6.7	468	US-11-102-240-150	Sequence 150, App
43	166	6.7	600	US-10-878-556A-43	Sequence 43, Appl
44	165.5	6.7	472	US-10-793-626-658	Sequence 658, App
45	164	6.6	495	US-10-821-234-1154	Sequence 1154, Ap

ALIGNMENTS

RESULT 1
US-10-878-556A-169
; Sequence 169, Application US/10878556A
; Publication No. US2005026399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; FILE OF INVENTION: HCV regulated protein expression
; TITLE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 169
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: humanpgp/chr12-q14221
; DATABASE ENTRY DATE: 2003-04-22
US-10-878-556A-169

Alignment Scores:

Pred. No.: 8.02e-09
Score: 280.50
Percent Similarity: 45.74%
Best Local Similarity: 20.92%
Query Match: 11.28%
DB: 6
Length: 1404
Matches: 118
Conservative: 140
Mismatches: 207
Indels: 99
Gaps: 21

US-09-837-344-41 (1-1482) x US-10-878-556A-169 (1-1404)

Oy 1 CAAAGAACAAACGATCTAGAACAGAGACGTCTTAAGAAAGTTGCAAGACAA 60
Db 740 GUCGlnAlaLeuGlnApeLeuGlnGlnAlaArgGlnLeuGlnThrApeLeuGlnLeuArg 759

QY	61	CAAAGCGATTAA-----GAAACAAGATAGACTTGCTTAAAGAAAAGTTGACAAAGACGACAAAGC	117
Db	760	AlaThrGluLeuSerLysGlnLeuGlnMetGluLysGlnLeuValSerSerThrArgLeu	779
QY	118	GATTATGAAACAAGAGAGAGACTTGCTTAAAGAAAAGTTGCACAGAACAAACAGCGATCTAGAA	177
Db	780	AspLeuGlnLysLysSerGlnValAlaLeuGlnUserLysLysGlnLysLysLeuThrLysGlnGlu	799
QY	178	CAAGAGAGACGT-----GCTAAAGAAAAGTTGCCAA	207
Db	800	GluGluLysLysValLeuLysGlnAspPheGlnThrLeuSerGlnGluThrLysValLeGln	819
QY	208	GAAACAACAA-----AGCGATTTAGAACAAAGACGAT	240
Db	820	HisGluGlnLeuLeuAsnAsnArgIleGlnThrThrValThrGluLeuGlnLysValLysMet	839
QY	241	GCTAAAGAAAAGTTGCCAAGAACAAACAAACGATTTAGAACAAAGTACACTTGCTTAA---	297
Db	840	GluLysGlnValAlaLeuMetThrGluLeuSerThrVal---LysAspLysLysSerLysVal	858
QY	298	---GAAAAGTTACAAAGACGCAAAAGCATTTAGACAAAG---	336
Db	859	SerAspSerLeuLysAsnSerLysSerGlnPheGlnLysGlnLysGlnLysGlnLysValAla	878
QY	337	-----AGACGTGCTTAAAGAAAAGTTGCCAAGAACAA---	366
Db	879	AlaIleLeuAspLeuGlnLysThrCysLysGlnLeuLysHisGlnLeuGlnValGlnMet	898
QY	367	CAAGCGATTTAGAACAAAGACAGACGCTGCTTAAAGAAAAGTTGCCAAGAACAAACAGCAT	426
Db	899	GluAsnThrLeuLysGlnLysGlnLysGlnLysLysSerLeuGlnLysGlnLysGlnLysValAla	918
QY	427	TTAGAACAAAGAGAGACTTGCTTAAAGAAAAGTTGCCAAGAACAA-----CAAGAC	474
Db	919	SerHisGlnLeuLysLeuGlnLysLeuAsnSerMetGlnGlnGlnLysGlnLysValAlaGlnAsn	938
QY	475	GATTTAGAACAAAGACAGACGCTGCTTAAAGAAAAGTTGCCAAGAACAAACAGCGATTTAGAA	534
Db	939	ThrLeuLysGlnAsnGlnLysGlnLysGlnGlnGlnGlnGlnGlnGlnGlnLysValAlaGlnLys	958
QY	535	CAAGAGAGACGCTTAAAGAAAAG---TTGCAGAACAAACAGACGATTTAGAA-----	585
Db	959	GlnSerSerGlnGlnLysLysLysGlnIleGlnValAlaLeuGlnGlnGlnLysLysValAla	978
QY	586	-----CAAGAGAGACGCTGCTTAAAGAAAAGTTGCCAAGACGACAAAGATTTAGAACAA	639
Db	979	ValLeuGlnLysThrGluLeuGlnLysAsnLysLysGlnGlnGlnLysLeuThrGlnAlaAlaGln	998
QY	640	AGGAAGCGCTGATCCGAAAAA-----AATTGGAAGAAAAAAG	678
Db	999	GluLeuAlaAlaGlnLysGlnLysValSerValLeuGlnAsnAsnThrGlnLysSerGln	1018
QY	679	GAAACATGAGATATATTATGACAGAGGATTTATGCTGCTTTAGAAATCCACGATATAGAA	738
Db	1019	GluThrThrPheLysGlnLeuGlnSerAspPheThrGlnArg---GlnSerGlu	1034
QY	739	CTTCATCATCAAAAATGAACGTCGATATATATATACCAATCAATCTTTACTTCACGAGAC	798
Db	1035	LeuLeuAlaThrArgGlnAspLeuLysSerValGlnGluLysLeuSerLeuValGlnGlu	1054
QY	799	AACAGAGGAATAGATGATTCACAGGAATATCTATATATAGAAAAAACAATATAGAGAA	858
Db	1055	AspLeuLysLeuSerAsnArgAsn---GlnIleGlnAsnGlnAsnLysLeuLysGlnGln	1072
QY	859	TCTATTACCAAAATGTTGAAGACGAAGCAAGATATATACATTAAGACATCTTGAA---GAA	915
Db	1073	LeuLysThrAlaLysAlaThrLeuGlnGlnAspSerAlaLysLysGlnGlnGlnGln	1092
QY	916	AAGAAAGATGCTTCAATAAACCAGAACAAAAGAAAGT-----AAATCGCTGACACTA	969
Db	1093	GluArgCysLysValAlaLeuGlnAspIleGlnLysGlnLysSerLeuLysGlnLysGlnLys	1112

[illegible]

Db 662 IleaaplleaglualapleuglualeuaspProaspheilemetlalelysgluvalglu 681
Qy 169 GATCTAGAACAGAGAGAGCT---GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 225
Db 682 Glnleugluylsgluylsglualeuglualeuglualeuglualeuglualeugluale 701
Qy 226 GAACAG 279
Db 702 AspyrRheglualagualaleuaglualeuglualeuglualeuglualeugluale 721
Qy 280 CAAGATGACTTGCTTAAGAAAAGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 324
Db 722 GluGlnRglllelyaspmetaspLeutrpgluGlnGlnGlnGlnGlnGlnGlnGlnGln 741
Qy 325 -----TTAGACAG 360
Db 742 MetGlnleugluayrgluylaleuGlnlulhlslysaenaArgmetSerArgmetleuglu 761
Qy 361 GAACACAAAGCGATTGTAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 762 AsparRaspLeupheValmetcArgleuylsalaatrglnSerValTyrGlnGlnGln 781
Qy 421 AGCGATTAGAACAG 471
Db 782 LeuylsglnRhegluGlnRglllelyaspmetaspLeutrpgluGlnGlnGlnGlnGln 801
Qy 472 AGCGATTAGAACAG 522
Db 802 ArgGlnRglllelyaspmetaspLeutrpgluGlnGlnGlnGlnGlnGlnGlnGlnGln 821
Qy 523 AGCGATTAGAACAG 582
Db 822 -----ArgRglllelyaspmetaspLeutrpgluGlnGlnGlnGlnGlnGlnGln 835
Qy 583 GAACAG 642
Db 836 GlnRglllelyaspmetaspLeutrpgluGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 853
Qy 643 AAGCGTAGCGAAAAA-----AATTTAGAAAAGAAAAGAAACATGAGAGATATTA 696
Db 854 -----Vallylelyleuglualeuglualeuglualeuglualeuglualeugluale 868
Qy 697 GCAGAGATTATATGCTGCTTTAGAAATACCAGTATAGAACTTCCATCAAGAAAATGA 756
Db 869 -----LeuGlnlulhlslysaenaArgmetSerArgmetleuglualeugluale 881
Qy 757 CGTGATATTAATATACATCAATCTTTACTCGAGACAAAGAGAGAGAGAGAGAGAGAG 816
Db 882 Arg-----ArgleuGlnlyaspSerleuSerArglyaspSerArgTrrpglyaspArg 899
Qy 817 GATTCGAAGAAATATCTAATATAGAAAAAACAATAGAGAAATCTATTACAACAATGTT 876
Db 900 AspSer-----GluGlnlyThrTrrpglyaspGlyProGlnlulhlsaspSer 913
Qy 877 GAAGAGAGAGAG-----GATATACATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 921
Db 914 GlnTrrpArgRglllelyaspmetaspLeutrpgluGlnGlnGlnGlnGlnGlnGlnGln 933
Qy 922 GATGCTCAATTAACAG 981
Db 934 ArgSerHlsaspRglllelyaspmetaspLeutrpgluGlnGlnGlnGlnGlnGlnGlnGln 952
Qy 982 TTAGAGAGAGTAATATTTCTGATGTTAATGATTTCAATTAAGATATGAGAGTAA 1041
Db 953 -----Gln 953
Qy 1042 ATAAGTCTGAATATGAG 1095
Db 954 ProSerleuayrgProaspaspaspArgValProayrgayrglymetaspaspArgGly 973
Qy 1096 -----TTAGAGAGAGATTTTAGAGCT 1113
Db 974 ProayrgayrglyProGlnlulhlsaspRglllelyaspmetaspArgRglllelyasp 993

Qy 1114 ATTGTGCATATGCAATTTCCAAAGAT-----GAAGAAAAC 1149
Db 994 -----SerTrrpArgSerHlsaspRglllelyaspmetaspLeutrpgluGlnGlnGln 1011
Qy 1150 ATAGAAATTAATTAAGAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1209
Db 1012 ArgGlnlySerHlsaspRglllelyaspmetaspArgRglllelyaspmetaspArg 1030
Qy 1210 GATGAGAGAGATTAAGAAAATCATCAGAAAGA 1239
Db 1031 AspArgGlySerTrrpArgHlsaspGln 1040

RESULT 5
US-10-878-556A-97
; Sequence 97, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 718
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: humangp/chr7-g9unh0
; DATABASE ENTRY DATE: 2003-01-07
US-10-878-556A-97

Alignment Scores:
Pred. No.: 2,71e-07 Length: 718
Score: 251.50 Matches: 116
Percent Similarity: 45.27% Conservative: 104
Best Local Similarity: 23.87% Mismatches: 175
Query Match: 10.12% Indels: 91
DB: Gaps: 23

US-09-837-344-41 (1-1482) x US-10-878-556A-97 (1-718)

Qy 1 CAAG 60
Db 44 GlnTrrpArgSerHlsaspRglllelyaspmetaspLeutrpgluGlnGlnGlnGlnGln 63
Qy 61 CAAG 120
Db 64 LeuileGlnValGlnayrgGlnlymetGlnmetGln---LeuSerHlsaspRglllelyasp 82
Qy 121 TTAG 180
Db 83 ValGlnleugluayrgalaleaserThrSeraleRglnSerHlsaspRglllelyaspArg 102
Qy 181 GAG 240
Db 103 AsnGlnGlnleuLeuThrArgGlnleuGlnleuGlnleuGlnleuGlnleuGlnleuGln 119
Qy 241 GCTAAAGAAAAGTTGCAAG 288
Db 120 AlaGlnGlnlymetGlnGlnGlnleuGlnleuGlnleuGlnleuGlnleuGlnleuGln 136
Qy 289 -----CTTGTAAGAAAAGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 336
Db 137 AsnleuaspHlsaspSerlyaspRglllelyaspmetaspLeutrpgluGlnGlnGlnGln 156
Qy 337 -----AGAGCTGCTTAAGAAAAGTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 384
Db 157 GlnThrileaspHlsaspRglllelyaspmetaspLeutrpgluGlnGlnGlnGlnGln 176
Qy 385 GAG 426

Db 680 ArgGlnbenLeuEngInLeuSerGluIleYsgInGlnIArgLYs----- 695
Qy 610 TTGCAGAGCAGCAAGAGATTTAGAACAAAGAAAGCTGATACGAAAAAATTAGAA 669
Db 695 ----- 695
Qy 670 AGAAAAAGAACATGAGATATATTAGACAGATTTATATGCGTTAGAAATACA 729
Db 696 ----- 703
Qy 730 GCTATAGAACTTCATCAGAAATGAACGCGATTTATATACCAATCAATCTTTTA 789
Db 704 ----- 716
Qy 790 CCTCAGCAACAAGAGGAAATAGTAGAGATTCCAAAGAAATATCTATTAATGAAAAACA 849
Db 717 SerIleGluAen----- 720
Qy 850 AATGAGAAATCTATTACAAATAATGTTGAAAGACGAAGGATATACATAAGACATCTT 909
Db 721 ----- 722
Qy 910 GAAGAAAAAGATGTGTTCAATAAACCCAGAACAAAGAAAGATTAATCTGTCACATA 969
Db 723 GlnApyIleYs----- 735
Qy 970 CAAATCATCATATAGACAGCAATAATTTCTGATGTTAATGATTTCAATAAGTAG 1039
Db 736 IleAenHleAleIleApyHleAen----- 750
Qy 1030 TATGAGCATGAAATTAAGTCTGATATATGACGATTCATTAAATAGATGAAGAAAGATGAT 1089
Db 751 ---ValGluAysAlaArgMetGluIlyrApy----- 761
Qy 1090 GAAGCTAGACGAATTTAGCGCTATGTGCAATATGACAAATTCCTCAAGATGAAGA--- 1146
Db 762 GlnThrIleApyAenLeuArgIleYsValIleYsLeuThrIleApyGluLeu 781
Qy 1147 ----- 1188
Db 782 GlyProValAenLeuAenAlaIleGluGlnPheGluLeuAenGluAlyrGlyThrPhe 801
Qy 1189 ---AATGAAATTTAGATGATTTAGATGACGAATGAAAAATCTCGAAGAAATTAATCT 1245
Db 802 LeuAenGluGlnArgThrApyLeuArgGluAlaYsGluThrLeuGluGlnIleIleHs 821
Qy 1246 GAAGAAAAATTAATAAGAAAGAAATATGAAAAACAAGGATTAATTTAAACCA 1305
Db 822 Glu----- 833
Qy 1306 AATGATAAAGTTTGTATGATGACGATTTAAAAAATATAAAATGATAGCAGGTTAAT 1365
Db 834 ThrPheHsAlaValGlnApyHsIlePheThrValPheYs----- 847
Qy 1366 AAGAAAAAGAAAAATTCATAAATCATTTGTTCTATATATTGACGAGACAAATGAATT 1425
Db 848 ----- 856
Qy 1426 TTACGATCGTGATGAG 1443
Db 857 LeuApyLeuThrGluApy 862

RESULT 7
US-11-074-176-134
; Sequence 134, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Kleenhammer, Todd R.
; APPLICANT: Rubsell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding

; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: PaSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-134

Alignment Scores:
Pred. No.: 2,97e-06 Length: 1189
Score: 230.50 Matches: 144
Percent Similarity: 40.21% Conservative: 129
Best local Similarity: 21.21% Mismatches: 191
Query Match: 9.27% Indels: 215
DB: Gaps: 33

US-09-837-344-41 (1-1482) x US-11-074-176-134 (1-1189)
Qy 1 CAAGAGACAAAGCGATCTAGAACAAAGAGAG-----CGTGCTAAAGAAAG 48
Db 175 GlnGlnAlaGlnAenGlnLeuYsYsThrGlnApyAenLeuIleApyIleAenApyLeu 194
Qy 49 TTGCAGAACAAACGAGATTTA-----GAACAGATAGACTTGCTAAAGAA 96
Db 195 ValYsGluLeuGlnSerApyGluGlnProLeuAenGlnGlnSerLeuAlaYsGlu 214
Qy 97 ---AAGTTACAGACGCAAGCGATTTAGAACAAAGAGACTTGCTTAA----- 144
Db 215 TyrIleApyGlnYsSerGlyLeuApyYsLeuYsSerLeuAlaPheGluIle 234
Qy 145 GAAGAAGTTGCAAGAACAAAGCGATCTGAACAAAG----- 183
Db 235 GlnAenIleAenGlnArgGluApyIleGlnYsSerAlaApyYsAenYsIleLeu 254
Qy 184 ---AGAGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCT 240
Db 255 LeuAlaYsLeuApyApyGluValYsApySerGlnAlaValThrGlnYsApyAla 274
Qy 241 GCTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGCTTAACTTAACTTAACT 297
Db 275 GlnYrYsYsYsLeuApyApyGluApyG---ApyHsThrGlnAenYsLeuYsLeu 293
Qy 298 ---GAAAGCTTACAAAGCGACAAAGCGATTTA-----GAACAAAGAGAGCGTCT 345
Db 294 SerYsApyLeuSerGlnLeuAenAlaSerLeuGlnMetAlaGlnGlnSerApyGlnPhe 313
Qy 346 AAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAG 405
Db 314 ApyApyAlaThrYsGlnGluYrYsApyAen-----GlnValYsGlnLeuYsGlnAen 331
Qy 406 TTGCAGAACAAACGATTTAGAACAAAGAGAGCTTGCTAAAGAAAGTTGCAAGAA 465
Db 332 LeuValGlnLeuYsAlaApyLeuApyGlnLeuYsYsGluYsYsYsLeuGlnApy 351
Qy 466 CAACAAAGCGATTTAGAACAAAGAGAG-----CGTGCTTAA 501
Db 352 GlnGlnApyValLeuYsIleGluApyGlnLeuThrGlnLeuAenGluApyPro 371
Qy 502 GAAAGTTGCAAGAACAAAGCGAT-----TTAGAACAA 537
Db 372 GlnGlnLeuAenYsYsYsApyApyIleApyAenYrYrMetGlnLeuGlnApy 391
Qy 538 GAGAGAGCTGCTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCT 597
Db 392 GlnAlaThrThrAenAenGlnIleValAenLeuAenSerApyLeuApyApySerGlnAla 411
Qy 598 -----GCTAAGAAAGTTGCAAGAGAGAGCAAGAGATTTA 633

```
Db 412 AspThrThrTyrGlnThrGlyAspValSerIysGlnLeuThrAspAlaGlnIlySerGlnLeu 431
Qy 634 GAACAAAGAGAGCGGTGATACGAAAAAATTAGAA--AGAAAAAGACATGGAGAT 690
Db 432 GluGlnLeuArgIleGluGlyLysIlySerLeuThrAspIlyAspGlnIlySerGlnIlyAsnIla 451
Qy 691 ATATTA----- 696
Db 452 IleValArgIleAsnIlySerGlnIlySerGlnLeuLeuThrGlnLeuThrAsnLeuArgGln 471
Qy 697 -----GCAGAGGATTTATATGCGTTAGAAATACAGCATATAGACTTCATCA 747
Db 472 ValIleAsnIleGlu-----ArgAsnGlnLeuGlnIlyValGlnIlaArgHis 487
Qy 748 GAA-----AATGAACGTGATTTATATACCATCATCTCTCT 786
Db 488 GluAlaLeuValAsnIleGlnIlySerArgHisGluGlyTyrTyrGlyValArgAsnVal 507
Qy 787 TTA----- 789
Db 508 LeuAsnHisLeuAsnAspPheAlaGlyValIleGlyAlaValGlyGlnLeuIleThrPhe 527
Qy 790 -----CCTGAGGACCAACAGA 804
Db 528 ProAlaGlnLeuGlnIlaIleAlaMetThrThrAlaLeuGlyGlyValGlnIlySerLeuIle 547
Qy 805 GGGATATGATGAGATCCAAAGAAATATCTATATATGAA--AAAACAAT----- 852
Db 548 ThrGluSerArgIleSerAlaArgAsnAlaIleAsnIlyLeuIlySerGlnIlySerGlnIly 567
Qy 853 -----AGAGATCTATT 864
Db 568 ArgAlaThrPheLeuProLeuAspGlyLeuArgGlnTyrGlyIleProGlnSerThrVal 587
Qy 865 ACACACA-----AATGTTGAAGCAGACGAGATATACATAAAGACATCTTGAAGAAAG 918
Db 588 ThrThrLeuIlySerIlyThrAspGlyPheArgGlyIleAlaSerAspLeuValGluSerIys 607
Qy 919 AAAGATGTTCAATTAACCAAGAAACAAAGAAATATCTGCTGACATCAAAATCAT 978
Db 608 ThrAsp-----GlnAspIleThrAlaIleIle--AsnTyr 618
Qy 979 ACATTGAGACAGATAATTTCTGAT----- 1005
Db 619 LeuLeuGlySerValIleValIleValAspThrIleAspThrAlaMetSerValAlaGlnArg 638
Qy 1006 GTTAATGATTTTCAATTAAGTAAGTATGAGATGAATGAATGCT----- 1050
Db 639 ValAsnAspGlyThrArgIleValIleThrLeuAspGlyAspValIleSerProGlyGlySerMet 658
Qy 1051 -----GAAATGACGATTCATTAATGATGAAAGAAAGATGATGAA 1092
Db 659 ThrGlyGlyGlnArgAsnGlnArgSerAsnSerProLeuGlnIlyThrAlaThrGlnIleAsn 678
Qy 1093 GACTTGAAC--GAATTTAAGCCTATTGTCATATGACATATTCGAAAGTGAAGAAAC 1149
Db 679 GlnLeuGlnIlySerGlnIleIlySerThrLeuIlySerGln-----AsnLeuAsnGlnIlyAspGlnAsp 696
Qy 1150 ATAGCAATTTATAAAGAACTAGAAAGTTGATGAGAAA-----AATGAAGAT 1197
Db 697 -----LysLeuGlnIlyAsnLeuValAspGlnSerAsnIlySerValAsnAlaGln 711
Qy 1198 TTATGATGATTTAGATGAAGAAATGAAATTCATCGAAGATTTATCTGAAGAAATTA 1257
Db 712 LeuGlnIlySerGlnIlySerAlaLeuArgGlnThrSerGlnIleAlaSerGlnIlyAlaIle 731
Qy 1258 AAA-----AAAGAAAGAAATATGAAGAAACAAAGATTAATTTAAACCAAT 1308
Db 732 SerPheGlnIlyGlnIlyGlnIlyValIlyValArgLeuLeuAspAlaAsn----- 747
Qy 1309 GATAAAGTTTGTATGATGACATATTTAA----- 1338
```

```
Db 748 -----ThrLeuTyrLysSerArgIleLysAspArgAsnAspArgIleGlnLeuLeuIys 765
Qy 1339 -----AAATATTAATAATGATTAAGCAGGT-----AATAAGAA 1371
Db 766 LysGlnIleLysGlnAlaAsnAspLysGlnMetLeuThrIlyGlnIlyGlnIlyGln 785
Qy 1372 AAGGAAATTCATTAATTCATTTGTTGATATATTGACGAGACAAATTTTACAG 1431
Db 786 LysAlaLysMetAsnAspLeuGlnIlySerIlySerIlySerAsnPheAsnIlySerGlnArg 805
Qy 1432 ATCTGATGATGATTTCTGAA-----GATATTAATAATTTATATGAAACTA 1479
Db 806 IleGlnAspGlnLeuSerIlySerLysLeuAspProIlyAlaValIlyThrAsnIlySer 824

RESULT 8
US-10-821-234-901
/ Sequence 901, Application US/10821234
/ Publication No. US20050255114A1
/ GENERAL INFORMATION:
/ APPLICANT: Labat, Ivan
/ APPLICANT: Stache-Crain, Birgit
/ APPLICANT: Andarmant, Susan
/ APPLICANT: Tang, Y. Tom
/ TITLE OF INVENTION: Methods for Diagnosis and Treatment of Presclampsia
/ FILE REFERENCE: 821A
/ CURRENT APPLICATION NUMBER: US/10/821,234
/ PRIOR FILING DATE: 2004-04-07
/ PRIOR APPLICATION NUMBER: US 60/462,047
/ NUMBER OF SEQ ID NOS: 1704
/ SOFTWARE: pc_seq_genes Version 1.0
/ SEQ ID NO 901
/ LENGTH: 1586
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-821-234-901

Alignment Scores:
Pred. No.: 3,816-06 Length: 1586
Score: 228.00 Matches: 126
Percent Similarity: 38.30% Conservative: 95
Best Local Similarity: 21.84% Mismatches: 168
Query Match: 9.17% Indels: 168
DB: Gaps: 23

US-09-837-344-41 (1-1482) x US-10-821-234-901 (1-1586)
Qy 1 CAAGAAACAACAAGCGATCTAGACACAGACAGCGTGAAGAAAG-----TTGCAG 54
Db 1025 GluGlnGlnArgLysAlaLeuGlnIlaLysValAlaAlaPheGlnIlySerGlnValIleGln 1044
Qy 55 GAACAAACAAGCGATTTTGAACAAGTACCTTGTGAAGAAAGTTACAAG----- 108
Db 1045 LeuGlnAlaSerHisArgIlySerGlnIlyAlaLeuGlnIlyArgLeuAspGlnValIleSer 1064
Qy 109 -----CAGCAAGCGATTTTGAACAAGACAGACTTGTGAAGAAAGTTG 153
Db 1065 ArgGlnLeuCyHisIleThrGlnIlySerHisAlaSerLeuArgAlaAspAlaGlnVal 1084
Qy 154 CAAGAAACAACAAGCGATTA-----GACAAAGAG 183
Db 1085 GlnGlnGlnGlnGlnGlnIleMetAlaGlnLeuHisSerIlyLeuGlnIlySerSerGlnIlyGln 1104
Qy 184 AGACGTGCTAA-----GAAGAATTGCAAGAAACAACAAGCGATTTGAACAAGACAGCGT 240
Db 1105 ValArgSerLysCySerGlnIlyLeuSerGlnIlyLeuHisGlyGlnLeuGlnIlyAlaArgAla 1124
Qy 241 GCTAAAGAAAGTTGCAAGACAAACAAGCGATTTGAA-----CAA 282
Db 1125 GluAsnSerGlnLeuThrGlnArgIleArgSerIleGlnIlyAlaLeuLeuGlnIlyGln 1144
Qy 283 GATAGACTTGTAAAGAAAGTTTACAAGACGACAAAGCGATTTGAACAAGACAGCGT 342
```


Db 1145 AlaArghepAlaGInaep---ValGInAlaSerGInAlaGInAlaepGInGInThr 1163
Qy 343 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAGAG-----AGA 390
Db 1164 ArgLeuYseGInLeuGInSerGInValSerGInLeuGInValAlaGInLeuArg 1183
Qy 391 CGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAA-----435
Db 1184 GluAlaValGInGInGInValValLeuAlaAlaSerGInLeuGInValSerLeuPLeuAla 1203
Qy 436 -----GAGAGCTTCTAAAGAAAAGTTGCAAGAA-----465
Db 1204 MecGInAlaLeuAlaThrAlaGInAlaLeuCyLeuGInLeuValSerLeuThrGIn 1223
Qy 465 -----465
Db 1224 AlaLeuGInLeuSerGInLeuGInLeuCyLeuLeuGInAlaGInThrMecGInAlaLeu 1243
Qy 466 -----CAACAAAGCGATTAGAACAGAGAGA 492
Db 1244 LeuAlaLeuLeuProGInLeuSerValLeuAlaGInGInAlaThrGInTrpLeuGIn 1263
Qy 493 CGTCTAAAGAAAAG-----TTGCAAGAACAA 522
Db 1264 AspLeuYseGInLeuGInProThrLeuLeuYseHisProProAlaProAlaGInProSer 1283
Qy 523 AGCGATTAGAACAGAGACGTCTAAAGAAAAGTTGCAAGAACAA---CAAAGCGAT 579
Db 1284 SerAspLeuAlaSerLeuArgGInAlaGInLeuTrpGInSerThrLeuGInAlaGIn 1303
Qy 580 TTAAGAACAGAGACGTCTAAAGAAAAGTTGCAAGAGCGAACAGATTAGAACAA 639
Db 1304 CysAspGInTrpArgSerLeuAlaGInLeuAlaGInLeuValGInLeuValGInLeu 1322
Qy 640 AGAAGCGTGTATACGAAAAAATTTAGAACAAAAAG-----678
Db 1323 LysSerValGInGInGInGInValTrpArgAlaLeuValGInAlaGInGInGIn 1342
Qy 679 -----GAACTGAGATATATTAGAACAGATTATATGCT 714
Db 1343 LeuGInLysSerArgValThrValLysHisLeuGInGInLeuValGInLysLeuLeuGIn 1362
Qy 715 CGTTAGAACAAATCCA-----GCTATAGAACTTCATCGAAGAAAGAA 756
Db 1363 GluLeuGInSerSerAspGInValArgGInHisThrLeuHisLeuGInAlaGInLeuGIn 1382
Qy 756 -----756
Db 1383 LysHisMetAlaAlaAlaSerAlaGInLysGInLeuTrpAlaLysGInValAlaGInLys 1402
Qy 757 CGTGGATATATTATACCAATCAATCTTTTACCCTCAGACAAACAGAGGAATAGTAGA 816
Db 1403 ArgGInLeuLeuLeuGInSerGInSerGInLeuAlaAlaLysSerGInAlaGInLys 1422
Qy 817 GATTCCAGAGAAATATCTTAATAGAAAAAAACAATAGAGAAATCTATTACAACAATGCT 876
Db 1423 GInSerAspGInLeuAlaLeuValArgGInLeuSerGInMet-----1437
Qy 877 GAAGAGCAAGAGATATACATTAAGGACATCTTGAGAAAGAAAGATGCTTACATA--- 933
Db 1438 -----LysSerHisValGIn-----AspGInAspHisAla 1447
Qy 934 -----AAACAGAACAAAGAAAGAAAGATTAATCTGTCGATCAAAATATCAT 978
Db 1448 GlyAlaProAlaSerSerProGInAlaProProAlaGInGInAlaProValGInLeuYs 1467
Qy 979 ACATTAGAGACAGTAAATATTTCTGATGTATATGTTTTCAAATAAGTAAGTAGAGAT 1038
Db 1468 ThrGInLeuGInTrpThrGInAlaLeuLeuLeuSerGInGInThrGInArg-----GIn 1485
Qy 1039 GAAATTAAGTGTGAATATGATGATTCATTAATAGATGAAGAAAGATGATGAAGACTTA 1098
Db 1486 LysLeuThrAlaGInUpheGInGInAlaGInThrSerAlaCyAspArgLeuGInGInGIn 1505

Qy 1099 GACGAATTTAAGCCTATTGTCATATGCAATATTTCCAGATGAGAAAGAAACATAGAAATT 1158
Db 1506 GInLysLeuArgThrAlaGInProLeuGInLysSerGInThrGInGIn-----1521
Qy 1159 TATTAAGAACTAGAGATTTGATAGAGAAAAATGAAAAATTTAGATAT-----1206
Db 1522 AlaSerGInLeuYseGInArgGInLeuGInLysGInLysLeuThrSerAspLeuGInArg 1541
Qy 1207 -----TTAGATGAGAAATAGAAAAATCATCGAAGAAATTTATCTGAAGAAAA 1254
Db 1542 AlaAlaThrArgLeuGInGInGInLeuLeuYseThrThrGInGInGInAlaArgGInLys 1561
Qy 1255 -----ATAAAAAA---GGAAGAAATATGAAAAAAACAAGATTAATAT 1296
Db 1562 AspThrValLysValSerGInGInGInGInLysGInLysGInLysAlaGInAspGInSer 1578

RESULT 9
US-10-878-556A-136
; Sequence 136, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 136
; LENGTH: 1410
; TYPE: PRF
; ORGANISM: Homo sapiens
; PUBLICATON INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/trbl_human
; DATABASE ENTRY DATE: 2003-02-28
US-10-878-556A-136

Alignment Scores:
Pred. No.: 6,99e-06 Length: 1410
Score: 223.00 Matches: 122
Percent Similarity: 39.13% Conservative: 103
Best local Similarity: 21.22% Mismatches: 186
Query Match: 8.97% Indels: 164
DB: Gaps: 23

US-09-837-344-41 (1-1482) x US-10-878-556A-136 (1-1410)

Qy 1 CAAGAACACAAAGCGATTAGAACAGAGACGTGCTAAAGAAAAG-----TTGCAA 54
Db 849 GInGInGInLysGInValAlaLeuGInValAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 868
Qy 55 GAACACAAAGCGATTAGAACAGATGAGCTTCTAAAGAAAAGTTACAAGAG-----108
Db 869 LeuGInAlaSerHisAspGInLysSerGInGInAlaLeuGInLysAspGInLysValSer 888
Qy 109 -----CAGCAAGCGATTAGAACACAGAGACGTGCTAAAGAAAAGTTG 153
Db 889 ArgGInLeuCyHisThrGInSerSerHisAlaSerLeuArgAlaAspAlaGInLysAla 908
Qy 154 CAAGAACACAAAGCGATCTA-----GAACAGAG 183
Db 909 GIn 928
Qy 184 AGACGTCTAA-----GAAAGTTGCAAGAACAAAGC 219
Db 929 ValArgSerLysCysGInGInLysSerGInLeuHisLeuGInSerGInSerGInAlaGIn 948
Qy 220 GATTAGAACAA-----GAG 234
Db 949 GInLysSerGInLeuThrGInArgGInLeuArgSerLeuGInAlaLeuLeuGInAlaGInGIn 968
Qy 235 AGACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGATGACTTGCT 294

[illegible]

QY	1288	GATATAATTAATTTTAAACCAATG-----ATAAAAGTTGTATGATGAG	1323
		: :	
Db	890	AspIyysIys--IleasnGIuMetSerIysGIuValThrIysLeuIysGIuAlaLeuIysnSe	909
QY	1330	CATA-----TTAAAAATTTAAAAAT	1350
		: :	
Db	909	rIeuserGIuIeuserIyrsertThrIrserserserIysArgIuIserGIuInIeugIuAl	929
QY	1351	GATTAAGCAGGTTTAATAAGAAAAAGAAAAATTCATAAAATCATTTGTTTCATATATTTGAC	1410
		: :	
Db	929	aleuGIuInGIuInValIysGIuInIeugIuInaInGIuInIeuaIaGIuCybeIysGIuInIsgI	949
QY	1411	CGAGACATGATAAATTTTACAGATGTCGATGAGCTAT	1447
		: :	
Db	949	ngIuValIleSerValIyTrArgMetHisIleuLeuTyR	961
RESULT 11			
US-10-821-234-1050			
; Sequence 1050, Application US/10821234			
; Publication No. US200502551141			
; GENERAL INFORMATION:			
; APPLICANT: Labat, Ivan			
; APPLICANT: Stache-Crain, Birgit			
; APPLICANT: Andarmani, Susan			
; APPLICANT: Tang, Y. Tom			
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia			
; FILE REFERENCE: 821A			
; CURRENT APPLICATION NUMBER: US/10/821,234			
; CURRENT FILING DATE: 2004-04-07			
; PRIOR APPLICATION NUMBER: US 60/462,047			
; PRIOR FILING DATE: 2003-04-07			
; NUMBER OF SEQ ID NOS: 1704			
; SOFTWARE: pc_seq_genes Version 1.0			
; SEQ ID NO 1050			
; LENGTH: 1410			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-821-234-1050			
Alignment Scores:			
Pred. No.: 1,69e-05 Length: 1410			
Score: 215.50 Matches: 85			
Percent Similarity: 48.82% Conservative: 60			
Best Local Similarity: 28.62% Mismatches: 107			
Query Match: 8.67% Indels: 45			
DB: 6 Gaps: 14			
US-09-837-344-41 (1-1482) x US-10-821-234-1050 (1-1410)			
QY	1	CAAGAACCAACAACGCGATTTCAGAACAGAGACGTCCTAAAGAAAGTTCAGAACAA	60
		: :	
Db	1086	GIuIysrGIuIysAlaIayGIuIeugGIuIuArgIuIuArgIysArgArg-----GIuGIuIn	1103
QY	61	CAAGCGCGATTTCAGAACAGATAGA-----	84
		: :	
Db	1104	GIuIuArgIysIeugIyGIuIeMetArgThrGIuIserIeIysnProIaIProPheserProIeu	1123
QY	85	CTTGCTTAAGAAAAAGTTACAGAGCAGCAGCAAGCGATTAGAA-----CAAGAGAGACTT	138
		: :	
Db	1124	ThrIaGIuInIeMetIysProGIuIuIysrProserThrIeugIuInArgProGIuInGIuThrVal	1143
QY	139	GCTAAAGAAAGTTCAGAACAGAACCAAGAGCGATCTAACAACAGAGAGACGTCGTAAGAA	198
		: :	
Db	1144	IleArgIuI--LeuGIuInProGIuInGIuInProArgThrIleGIuIuArgIysrIeugIuIn	1162
QY	199	AAGTTGCAAGAACCAAGAGCGATTTCAGAACAA-----	231
		: :	
Db	1163	TyrIleThrValserIysrGIuGIuIeIeuserSerGIyThrValCysProProAepProTyr	1182
QY	232	GAGAGACGTCGTTAAAGAAAGTTCAGAACAGAACCAAGAGCGATTTCAGAACAGATGACTT	291
		: :	
Db	1183	LysrArgAspAlaIayrGIuIysrIeugIuIeGIuInIeGIuInIeMetHisIleValaIePmetIeu	1202

Mon Dec 12 08:09:37 2005

us-09-837-344-41.rapbn

Page 16

Search completed: December 7, 2005, 07:16:34
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:30:46 ; Search time 5601.15 Seconds

(without alignments)
2639.584 Million cell updates/sec

Title: US-09-837-344-31

Perfect score: 1563
Sequence: 1 SDLEQRRAKEKLEQEQSDL.....GHLEKKDGSIRPEQKEDKS 316

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xih
-O/cgnt2_1/USPTO.spool/h/US09837344/runat.05122005.133532.750/app.query.fasta_1.718
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09837344 @CGN 1 1 6731 @runat.05122005.133532.750 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUDRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	396	25.3	317	10	T09652 0224m3 gmbp
2	390.5	25.0	777	10	BX200721 Danilo rer
3	350.5	22.4	619	8	CX031474 1343594 N
4	345.5	22.1	848	9	A2672705 ENTMB16TF
5	328.5	21.0	743	10	C2247380 AIPA-aaef3
6	325	20.8	718	9	A2972907 2M0246M24
7	322.5	20.6	878	9	BH149306 ENTMB89TF

8	320.5	20.5	1261	2	BG533805	BG533805 602562478
9	320	20.5	894	9	AZ688399	AZ688399 ENTMB244TF
10	319.5	20.4	868	9	AZ563100	AZ563100 ENTMB015TF
11	319	20.4	771	7	CN036384	CN036384 nm.15.c6
12	317.5	20.3	650	9	AZ337339	AZ337339 IM0068B03
13	316.5	20.2	1041	10	CW923205	CW923205 EDCAX27TR
14	316	20.2	612	7	CN508745	CN508745 AGENCOURT
15	314	20.1	743	3	BJ395948	BJ395948 BJ395948
16	308.5	19.7	814	9	AZ672545	AZ672545 ENTMB44TF
17	307.5	19.7	1327	10	AG429715	AG429715 Mus muscu
18	305	19.5	1264	2	BG541846	BG541846 602569739
19	304.5	19.5	1723	10	AG330644	AG330644 Mus muscu
20	304	19.4	871	9	AZ530343	AZ530343 ENTMB91TF
21	303	19.4	704	7	CF864210	CF864210 ps2S009XI
22	303	19.4	1794	10	AG382218	AG382218 Mus muscu
23	301.5	19.3	604	3	BQ264044	BQ264044 faa21e04.
24	300	19.2	1205	10	AG278065	AG278065 Mus muscu
25	299	19.1	890	9	BH163382	BH163382 ENTMB21TF
26	299	19.1	1465	10	AG386901	AG386901 Mus muscu
27	297.5	19.0	737	1	AJ724363	AJ724363 AJ724363
28	297.5	19.0	1921	10	AG365436	AG365436 Mus muscu
29	297	19.0	693	7	CK711667	CK711667 ZP201-P00
30	297	19.0	1945	10	AG370924	AG370924 Mus muscu
31	296.5	19.0	1567	10	AG349825	AG349825 Mus muscu
32	295	18.9	670	1	AL725110	AL725110 AL725110
33	295	18.9	1439	10	CG743836	CG743836 P036-1-H0
34	294.5	18.8	1957	10	AG382393	AG382393 Mus muscu
35	294	18.8	1933	4	AY325173	AY325173 Rattus no
36	293.5	18.8	548	11	DR2M2T	AL735260 Danilo rer
37	293.5	18.7	1409	10	AG390910	AG390910 Mus muscu
38	293	18.7	1969	10	AG330667	AG330667 Mus muscu
39	290	18.6	1425	10	AG310594	AG310594 Mus muscu
40	289.5	18.5	869	7	CK159167	CK159167 FGAS04056
41	289.5	18.5	1533	10	AG335255	AG335255 Mus muscu
42	289.5	18.5	1685	10	AG390959	AG390959 Mus muscu
43	289	18.5	1596	10	AG430370	AG430370 Mus muscu
44	288	18.4	787	3	BJ375707	BJ375707 BJ375707
45	287.5	18.4	681	3	BJ345323	BJ345323 BJ345323

ALIGNMENTS

RESULT 1

LOCUS

T09652 317 bp DNA linear GSS 05-DEC-2000

DEFINITION 0224m3 gmbpFHB3.1, G. Roman Reddy Plasmodium falciparum genomic

T09652
T09652.1 GI:319484ACCESSION
T09652
KEYWORDS
GSS.
Plasmodium falciparum (malaria parasite P. falciparum)SOURCE
Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.REFERENCE
1 (bases 1 to 317)
Reddy,G.R., Chakrabarti,D., Schuster,S.M., Ferl,R.J., Almira,E.C.AUTHORS
and Dame,J.B.TITLE
Gene sequence tags from Plasmodium falciparum genomic DNA fragmentsJOURNAL
Proc. Natl. Acad. Sci. U.S.A. 90, 9867-9871 (1993)COMMENT
8234327
Other GSSes: 0224m7Contact: Dame JB
Dept. of Pathobiology, College of Veterinary MedicineUniversity of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USATel: 352 392 4700
Fax: 352 392 9704Email: damej@mail.ufl.edu
Seq Primer: T3Class: Shotgun.
Location/Qualifiers
1..317
/organism="Plasmodium falciparum"

Alignment Scores:
Prod. No.: 8 73e-26 Length: 619
Score: 350.50 Matches: 73
Percent Similarity: 64.73% Conservative: 61
Best Local Similarity: 35.27% Mismatches: 70
Query Match: 22.42% Indels: 3
Gaps: 8

US-09-837-344-31 (1-316) x CX031474 (1-619)

ORIGIN

FEATURES
source
1. 619
/organism="Oncoerythrus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/lab_host="DH108"
/clone_lib="NCCWA 10RT#3"
/note="Vector: PCMV Sport6.0; WTV oocyte library RT-EGG."

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus. 1 (bases 1 to 619)
Yao, J., Gahr, S. and Rexroad, C.E.
10RT#3 egg, NCCWA/MVU ESP Project, Phase II
Unpublished (2004)
Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Leetown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 x2129
Fax: 304 725 0351
Email: crexroad@nccwa.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross_match v0.990329.
Plate: 120 row: 1, column: 8
Seq primer: GTAATAGCACTCACTATAGGG.
Location/Qualifiers

7 ArGArGAlaLySgIuLyLeuGInGInGIn--SerApLeuGInGInSerApLeu 25
1 CGGAGATGAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 60
26 AlAlaLySgIuLyLeuGInGInGInSerApLeuGInGInGInGInGInGIn 45
61 GAAAGGAGAAAGCCCTTGAGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAG 120
46 LyLeuGInGInGInGInSerApLeuGInGInGInGInGInGInGInGInGIn 65
121 AGAGCCCTTAGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCTTG 180
66 GluGInGInSerApLeuGInGInGInGInGInGInGInGInGInGInGInGIn 85
181 GAAACAGAGAAAGCCCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
86 SerApLeuGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 105
241 AAGCCCTTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCTTG 300
106 GluGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 125
301 CACAG 360
126 ArGArGAlaLySgIuLyLeuGInGInGInGInGInGInGInGInGInGInGIn 145
361 AAAAG 420
146 LySgIuLyLeuGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 165
421 AAG 480
166 LeuGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 185
481 GCTTGGAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCTTGGA 540

Alignment Scores:
Prod. No.: 4.12e-25 Length: 848
Score: 345.50 Matches: 91
Percent Similarity: 62.21% Conservative: 72
Best Local Similarity: 34.73% Mismatches: 80
Query Match: 22.10% Indels: 19
Gaps: 7

US-09-837-344-31 (1-316) x AZ672705 (1-848)

ORIGIN

FEATURES
source
1. 848
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOSt; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, edn. M. Vaubin and B. Barrell, Oxford University Press, 1999)."

RESULT 4
AZ672705/c 848 bp DNA linear GSS 14-DEC-2000
LOCUS
DEFINITION
Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.
ACCESSION
AZ672705
VERSION
AZ672705.1 GI:11809851
KEYWORDS
GSS.
SOURCE
Entamoeba histolytica
ORGANISM
Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE
1 (bases 1 to 848)
Loftus, B., Van Aken, S. and Fraser, C.
Determination of clone end sequences from Entamoeba histolytica
HM1:IMSS sheared DNA library
Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 26
High quality sequence stop: 739.
Location/Qualifiers

186 GluGInSerApLeuGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 205
541 CAG 597
206 AspLeuGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 212
598 ---ATCGAAAG 615

5 GluGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 13
773 CAG 714

QY 14 GINGINGInSerAspLeuGInGInaPArGLeuAlaLySGLuLySLeuGInGIn 33
Db 713 CAAGAGAGGCTATTGCAATTGAAGAGAAAGCTTCAAAAGAAATTGACAGAGAGAA 654
QY 34 GInSerAspLeuGInGInGInuArGArGAlaLySGLuLySLeuGInGInSerAsp 53
Db 653 CCAAAAGCCCAAGAAAGAGATGAAGGTTAAAGAA-----GAGAGAGAGAGATTAGC 600
QY 54 LeuGInGInGInuArGArGAlaLySGLuLySLeuGInGInSerAspLeuGIn 72
Db 539 TTAGAGAGCGAAGCACTCCAAAAGAAATTGAGAAAGAAAGAAAGCAAAAGAGAA 540
QY 73 GINGInuArGArGAlaLySGLuLySLeuGInGInSerAspLeuGInGInaPArG 92
Db 539 GAGAAAGAAAGCAAAAGAGAA-----GAGAAAGAAAGCAAAAGAGAGAGAA 483
QY 93 LeuAlaLySGLuLySLeuGInGInSerAspLeuGInGInuArGArGAlaLyS 112
Db 482 CAAGCAAAAGAGAGAA---GAGAAAGCAAGCAAAAGAGAGAAAGAAAGCAAGAA 426
QY 113 GInuLySLeuGInGInGInSerAspLeuGInGInuArGArGAlaLySGLuLySLeu 132
Db 425 GAAAGAA---CGAAGAAAGAAAGCAAGAGAAAGCAAGAAAGAAAGCAAAAGACTGAA 369
QY 133 GINGInuGInSerAspLeuGInGInuArGArGAlaLySGLuLySLeuGInGIn 152
Db 368 GAGAGAGAGAAATTAAATTAGAGAGAAAGAAAGCAAAAGAGAGAAAGAGAAAGAA 309
QY 153 GInSerAspLeuGInGInGInuArGArGAlaLySGLuLySLeuGInGInSerAsp 172
Db 308 GCAAAAGAACTTGAAGAGAGAAAGAAAGCAAAAGAACTTGAAGAGAGAGAAATTAA 249
QY 173 LeuGInGInGInuArGArGAlaLySGLuLySLeuGInGInSerAspLeuGIn 192
Db 248 TTAGAGAGAAAGAAAGCTTGAAGAAAGAAAGCAAGAAAGAAAGAAAGAAAGAA 189
QY 193 GInuArGArGAlaLySGLuLySLeuGInGInGInuArGArGAlaLySGLuLySLeu 212
Db 188 GAGAGAGAGTTAAATTAAAGAGAGAGAAAGCTTCAAAAGAACTTGAAGAGAGAGAA 129
QY 213 AArPThLySGLuLySLeuGInGInuArGArGAlaLySGLuLySLeuGInGInSerAsp 232
Db 128 GAGAGAAAGAGAGAGAGCTTAAATATGATTTATGATGAGCTTACCTGAAAGCCGA 69
QY 233 TArGInuArGLeuGInuLeuPArGAlaLySGLuLySLeuGInGInSerAspLeuGIn 252
Db 68 AATGGGCTAGAGAAATTAAAGCTAGAGAAAGAGCTCAGTGAAGAGAGAT-----TAT 15
QY 253 ILePArG 254
Db 14 CTTCCA 9

RESULT 5
CZ247380 743 bp DNA linear GSS 11-FEB-2005
LOCUS A1AA-aat35a12.b1 Ancylosetoma caninum whole genome shotgun library
DEFINITION (A1AA)GSS 001) Ancylosetoma caninum genomic, genomic survey sequence.
ACCESSION CZ247380
VERSION CZ247380.1 GI:59632821
KEYWORDS GSS.
SOURCE Ancylosetoma caninum (dog hookworm)
ORGANISM Ancylosetoma caninum
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
Ancylostomatoidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
1 (bases 1 to 743)
Mitreva, M., McCarter, J.P., Pape, D., Rittler, E., Tsagaris, V.I., R.,
Ronko, I., Martin, J., Wylie, T., Dante, M., Meyer, R., Messina, D.,
Waterson, R.H., Clifton, S.W. and Wilson, R.
Genome Survey sequences from the parasitic nematode Ancylostoma
caninum
JOURNAL Unpublished (2004)
COMMENT Contact: Mitreva M

Washington University in St. Louis
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: nematode@watson.wustl.edu
Genomic DNA provided by John Hawdon (mcmj@wumc.edu) DNA
sequenced by Washington University Genome Sequencing Center
Class: shotgun.

FEATURES
source
location/Qualifiers
1..743
/organism="Ancylostoma caninum"
/mol_type="genomic DNA"
/strain="Baltimore"
/db_xref="taxon:29170"
/dev_stage="Adult"
/lab_host="GS10"
/clone_lib="Ancylostoma caninum whole genome shotgun
library (A1AA)GSS 001"
/note="Vector: pOTW13; Site 1: BstXI; Site 2: BstXI;
Ancylostoma caninum genomic DNA was randomly sheared,
end-repaired and size fractionated to enrich for 2-4 kb
fragments. Genomic DNA was provided by John Hawdon
(mcmj@wumc.edu) at George Washington University.
Sequencing by Washington University Genome Sequencing
Center, St. Louis, MO."

ORIGIN
Alignment Scores:
Pred. No.: 1,926-23 Length: 743
Score: 328.50 Matches: 59
Percent Similarity: 65.32% Conservative: 86
Best Local Similarity: 26.58% Mismatches: 74
Query Match: 21.02% Indels: 3
DB: 10 Gaps: 2

US-09-837-344-31 (1-316) x CZ247380 (1-743)

QY 4 GINGINGInuArGArGAlaLySGLuLySLeuGInGInGInSerAspLeuGInaPArG 23
Db 17 CAACAGCAACAAACGACAAACAAACAAACAAACGTTAAACAAACAGCAACAAACAA 76
QY 24 ArGLeuAlaLySGLuLySLeuGInGInGInGInSerAspLeuGInGInuArG 41
Db 77 CGACAAACGACAAACAAACCTTCAACAAACGACAAACAAACAAACAAACAAACGA 136
QY 42 ArGAlaLySGLuLySLeuGInGInGInSerAspLeuGInGInuArGArGAlaLyS 61
Db 137 CGACAAACGACAAACAAACGACAAACAAACAAACAAACAAACAAACAAACAA 196
QY 62 GInuLySLeuGInGInGInSerAspLeuGInGInuArGArGAlaLySGLuLySLeu 81
Db 197 CAACAGCAACAAACGACAAACAAACAAACAAACAAACGACGACGACAAACGACAA 256
QY 82 GINGINGInSerAspLeuGInGInaPArGLeuAlaLySGLuLySLeuGInGIn 101
Db 257 CAACAGCAACAAACAAACAAACAAACGACAAACAAACGACAAACAAACGACGA 316
QY 102 GInSerAspLeuGInGInGInuArGArGAlaLySGLuLySLeuGInGInSerAsp 121
Db 317 CGACAAACGACAAACGACAAACAAACAAACAAACAAACAAACAAACAAACAA 376
QY 122 LeuGInGInGInuArGArGAlaLySGLuLySLeuGInGInGInSerAspLeuGIn 141
Db 377 CGACAAACAAACAAACGACAAACGACAAACGACAAACGACAAACGACAAACGA 436
QY 142 GInuArGArGAlaLySGLuLySLeuGInGInGInSerAspLeuGInGInuArGArG 161
Db 437 CAACAGCAACGACAAACAAACAAACGACAAACGACAAACGACAAACGACAA 496
QY 162 AlaLySGLuLySLeuGInGInGInSerAspLeuGInGInuArGArGAlaLySGLu 181
Db 497 CAACAAACGACAAACGACAAACAAACAAACAAACGACAAACGACAAACGACAA 556

Oy		182	LysleuglnlgluglnInSerAspleuglnlglInUArGAGAlAlySGlLySeugln	201
Db		557	CGACTACGACAACAACGACGA---CAACAACTGTGGACGACTTACACAAACGACAA	613
Oy		202	GluGlnglnAASpLeuglnUGlnArGLySaIaaSpThrLylSAsnLeuglnArGLyS	221
Db		614	CACACAGCAGCAACTAACGACGAACTACTACAAAGCAAACAATTAACGACAA	673
Oy		222	Lysgltu 223	
Db		674	CGACAA 679	
RESULT 6				
LOCUS	AZ972907	718 bp	DNA	linear GSS 27-APR-2001
DEFINITION	2M0246W24R Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UGC2M0246W24 R, genomic survey sequence.			
ACCESSION	AZ972907			
VERSION	AZ972907.1	GI:13844134		
KEYWORDS	GSS.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	1 (bases 1 to 718) Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacez,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R. Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts Unpublished (2000) Contact: Robert B. Weiss University of Utah Genome Center Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA Tel.: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0246 row: M column: 24 Seq primer: CACACAGGAACAGCTATGACC Class: plasmid ends High quality sequence stop: 718. Location/Qualifiers 1..718 /organism="Mus musculus" /mol_type="genomic DNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUGC2M0246W24" /sex="Female" /lab_host="E. coli strain XL10-Gold, T1-resistant, F-" /clone_1lb="Mouse 10kb plasmid UUGC2M library" /note="Vector: pMD42uv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi 4732114 gb AF129072.1), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to			
JOURNAL COMMENT				
TITLE				
FEATURES				
SOURCE				

adapto red vector DNA, and transformed into chemically-competent *E. coli* X10-Gold (Stratagene) cells and selected for ampicillin resistance."

Alignment Scores:

Pred. No.:	4.2e-23	Length:	718
Score:	325.00	Matches:	75
Percent Similarity:	63.41%	Conservative:	55
Best Local Similarity:	36.59%	Mismatches:	53
Query Match:	20.79%	Indels:	22
DB:	9	Gaps:	5

US-09-837-344-31 (1-316) x AZ972907 (1-718)

Oy		4	GluGlnGluAArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnAsp	23					
Db		706	CAGAGAGACGAGCAGCAGAGAACACACAGCAGAGCAGAGCAGAGCAGAGCAGAG	647					
Oy		24	ArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgAla	43					
Db		646	CAGCAGCAGAGAGAGCAGCTACGAGACAGCAGCAGAGAGCAGAGCGAAAG-----	596					
Oy		44	LysGluLysLeuGlnGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnArgAlaLysGluLys	63					
Db		595	CAGAGACGAGCAGCAGAGCAGAGCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	536					
Oy		64	LeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnLysArgAlaLysGluLysLeuGlnGln	83					
Db		535	GAGCAGAGACGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAG	476					
Oy		84	GlnGlnSerAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSer	103					
Db		475	CAGCAGCAGAGCAGCAGCAG-----CAGCAGCAGAGCAG	440					
Oy		104	AspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	123					
Db		439	GAGCAGAGCAGCAGAGCAGCAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAG	380					
Oy		124	GlnGlnLysArgAlaLysGluLysLeuGlnGlnGlnGlnGlnSerAspLeuGlnGlnGln	142					
Db		379	CAGAGAGCAGCAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAG	320					
Oy		143	ArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnGln	162					
Db		319	CAGCAGCAGCAGCAGCAGCAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAG-----	269					
Oy		163	LysGluLysLeuGlnGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnGlnGln	182					
Db		268	-----GAGAGCAGACACAGCAGCAGAGCAGAGCAGAG-----CAGAGCAG	230					
Oy		183	LeuGlnGlnGlnGlnGlnSerAspLeuGlnGlnGlnLysArgAlaLysGluLysLeuGlnGln	202					
Db		229	GAGCAGAGCAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGATTGTATGA	170					
Oy		203	GlnGlnArgAspLeu	207					
Db		169	TTCCAATGTCAGCTT	155					
RESULT 7									
BH149306/c									
LOCUS	BH149306	878 bp	DNA	linear GSS 28-AUG-2001					
DEFINITION	ENTRPH8PTF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.								
ACCESSION	BH149306								
VERSION	BH149306.1	GI:15321785							
KEYWORDS	GSS.								
SOURCE	Entamoeba histolytica								
ORGANISM	Eukaryota; Eutrombicidae; Entamoeba.								
REFERENCE	1 (bases 1 to 878) Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.								
AUTHORS	Determination of clone end sequences from Entamoeba histolytica								
TITLE									


```
Db      15 AGAGAGCGCTGAAAGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGCT 74
Qy      27 LygGluLybLeu---GlnGluGlnGlnSerAapLeuGlnGlnGlnGlnAArgAlaLybGlu 45
Db      75 AAAAGAAAAATTGAAACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 134
Qy      46 LybLeuGlnGlnGlnSerAapLeuGlnGlnGlnGlnAArgAlaLybGluLybLeuGln 65
Db      135 -----GAAAGAAACAAAGAAAGCGCTTGAAGAGAG-----AAAAATTAATTGAA 179
Qy      66 GlnGlnGlnSerAapLeuGlnGlnGlnGlnAArgAlaLybGluLybLeuGlnGlnGln 85
Db      180 GAAAGAAACAAAGAAAGATTAGAGAGAAAGAAAGAAAG-----AAGGACAGAAAGAAAGAA 230
Qy      86 SerAapLeuGlnGlnAapArgLeuAlaLybGluLybLeuGlnGlnGlnGlnSerAapLeu 105
Db      231 AAACGCTTGAAGCGTCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 287
Qy      106 GlnGlnGlnAArgAlaLybGlu-----LybLeuGlnGlnGlnGlnGlnSer 120
Db      288 GAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCGTTAGAGAAAGAAAGAAAGAAAGCGTTAG 347
Qy      121 AapLeuGlnGlnGlnGlnAArgAlaLybGluLybLeuGlnGlnGlnGlnSerAapLeuGln 140
Db      348 AGAAAGAAAGAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 407
Qy      141 GlnGlnGlnAArgAlaLybGlu-----LybLeuGlnGlnGlnGlnGlnSerAap 155
Db      408 GAAAGCGCTTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCGTTAA 467
Qy      156 LeuGlnGlnGlnGlnAArgAlaLybGluLybLeuGlnGlnGlnGlnSerAapLeuGln 175
Db      468 GCTGAGCGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 524
Qy      176 GlnAArgAlaLybGlu-----LybLeuGlnGlnGlnGlnGlnSerAapLeuGln 191
Db      525 GAAAGCTTGAAGAAAGAAAGAAAGAAAGAAAGCGTTAAAGCGGAAAGAAAGAAAGAAAGAG 584
Qy      192 GlnGlnGlnAArgAlaLybGluLybLeuGlnGlnGlnGlnAArgAapLeuGlnGlnAArg 211
Db      585 GAGGAGCTTGAAGAAAGAAAGAAAGCGCAAGAAAGAAAGAAAGAAAGCTTGAAGAAAGAAAG 644
Qy      212 AlaAapThrLybLeuAapLeuGlnAArgLybLeuGlnGlnGlnGlnGlnGlnGlnGln 231
Db      645 AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 704
Qy      232 LeuTyGlnAArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 251
Db      705 GAAAAAAAGCGT-----AAAGCGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 740
Qy      252 TyrTleProHiseGlnSerAapLeuProGlnAapArgAlaGlnAapSerAapSerAap 271
Db      741 AAAGAAAGAAAGAAAGCGGCGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 800
Qy      272 GlnTleSerIleIleGlu-----LybThrAapArgGlnGlnGlnGlnGlnGlnGlnGln 286
Db      801 GAAAGCTTGAAGTACAGATGAAAGAAAGAAAGCGTTAGAGAAAGAAAGAAAGAAAGCGTTAG 860
Qy      287 AapAlaGlnGlnAArgAArgAlaLybGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 303
Db      861 TAAAGCAAGAA-----GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 881

RESULT 10
AZS36100      868 bp      DNA      linear      GSS 03-NOV-2000
LOCUS      AZS36100
DEFINITION      ENT00157F Entamoeba histolytica Shared DNA Entamoeba histolytica
ACCESSION      AZS36100
VERSION      AZS36100.1 GI:11093047
KEYWORDS      GSS.
SOURCE      Entamoeba histolytica
ORGANISM      Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
```

```
REFERENCE      1 (bases 1 to 868)
AUTHORS      Loftus,B., Van Aken,S. and Fraser,C.
TITLE      Determination of clone end sequences from Entamoeba histolytica
JOURNAL      HM1:IMSS sheared DNA library
COMMENT      Unpublished (2000)
              Contact: Brendan J Loftus
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0208
              Fax: 301 838 3543
              Email: b.loftus@tigr.org
              Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
              DNA library
              Seq primer: M13-Forward
              Class: shotgun
              High quality sequence start: 74
              High quality sequence stop: 837.
FEATURES
source
1..868
location/Qualifiers
1..868
1..868
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHOSt1; Site 1: Bat I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaubin and B.
Bartell, Oxford University Press, 1999)."
```

ALIGNMENT SCORES:

Pred. No.:	1..926-22	Length:	868
Score:	319.50	Matches:	93
Percent Similarity:	52.81%	Conservative:	76
Best Local Similarity:	29.06%	Mismatches:	88
Query Match:	20.44%	Indels:	63
DB:	9	Gaps:	13

US-09-837-344-31 (1-316) x AZS36100 (1-868)

```
Qy      4 GlnGlnGlnAArgAlaLybGluLybLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 23
Db      29 GAGGCTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 82
Qy      24 ArgLeuAlaLybGluLybLeu---GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 42
Db      83 -----GCTAAAGAAAGAAATTTGACAGAAAGAAAGAAAGAAAGAAATTTGAGAGAAAG 136
Qy      43 AlaLybGluLybLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 62
Db      137 AAAAAAGAA-----GAAAGAAAGAAAGAAAGCGCTTGAAGAGAG-----AAAAAA 181
Qy      63 LybLeuGlnGlnGlnGlnSerAapLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 82
Db      182 TTATTGGAAGAAAGAAAGAAAGATTAGAGAAAGAAAGAAAG-----AAAGGACAGAA 232
Qy      83 GlnGlnGlnSerAapLeuGlnGlnGlnGlnAapArgLeuAlaLybGluLybLeuGlnGlnGln 102
Db      233 GAAAGAAAGAAAGCGTTGAACTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 289
Qy      103 SerAapLeuGlnGlnGlnGlnAArgAlaLybGlu-----LybLeuGlnGln 117
Db      290 AGAAAGAAAGAGAGAAAGAAAGAAAGAAAGAAAGCGTTAGAGAAAGAAAGAAAGAAAGAA 349
```


QY 118 GlnGlnSerAsePLeuGlnGlnGlnArgAlaIysGlnIleuGlnGlnSer 137
DB 350 GAACGTAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 409
QY 138 AsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 152
DB 410 TTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 469
QY 153 GlnSerAsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 172
DB 470 AAACGTAAAGCTGAGCGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 526
QY 173 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 188
DB 527 GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 586
QY 189 AsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 208
DB 587 AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 646
QY 209 GlnArgIleuAsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 228
DB 647 GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 694
QY 229 AlaGlnAsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 248
DB 695 -----GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 706
QY 249 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 268
DB 707 CGA-----CCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 736
QY 269 AsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 288
DB 737 GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 796
QY 289 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 304
DB 797 ACCCGGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 856

RESULT 11
CN036384 771 bp mRNA linear EST 29-MAR-2004
LOCUS CN036384
DEFINITION nm.15_c6_t3_Mach Ambystoma mexicanum cDNA, mRNA sequence.
ACCESSION CN036384.1 GI:45806755
VERSION
KEYWORDS EST.
SOURCE Ambystoma mexicanum (axolotl)
ORGANISM Ambystoma mexicanum
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
Ambystoma.
1 (bases 1 to 771)
Putte, S., Smith, J.J., Walker, J.A., Rondet, M., Westrock, D.,
Morgan, J., Samuels, A.K., Kump, K., King, D.C., Marrese, N.J.,
Haberman, B., Tanaka, E., Bryant, S.V., Gardiner, D.M., Parichy, D.M.
and Voss, S.R.
From biomedicine to natural history research: EST resources for
Ambystoma salamanders
BMC Genomics 5 (1), 54 (2004)
15310388
Contact: SR Voss
Department of Biology
University of Kentucky
TH Morgan Building, Lexington, KY 40506, USA
Tel: 859 257 9888
Fax: 859 257 1717
Email: sryos@uky.edu
The EST is quality trimmed at the ends with a 20 base window and
quality threshold of 15 (phred quality score). Please visit
http://salamander.uky.edu for any information (trace, quality files
etc) regarding this EST.

FEATURES
source Location/Qualifiers
1..771
/organism="Ambystoma mexicanum"
/mol_type="mRNA"
/db_xref="taxon:8296"
/tissue_type="limb Blastema and proximal limb tissue
collected from larvae on days 1-6 of regeneration"
/clone_lib="Mach"

ORIGIN
Alignment Scores:
Pred. No.: 1,87e-22 Length: 771
Score: 319.00 Matches: 74
Percent Similarity: 66.09% Conservative: 78
Best Local Similarity: 32.17% Mismatches: 66
Query Match: 20.41% Indels: 12
DB: 7 Gaps: 8

US-09-837-344-31 (1-316) x CN036384 (1-771)

QY 2 AsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 21
DB 89 GATTGTGAGCGTGAAGATAGACCGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 148
QY 22 GlnAsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
DB 149 GATGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 208
QY 41 ArgArgAlaIysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
DB 209 CGT-----GACGGGTCCGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 256
QY 61 IysGlnIleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
DB 257 AGAGAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 316
QY 81 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
DB 317 GAGCGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 373
QY 101 GlnGlnSerAsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 117
DB 374 AAGAGCGGAGAGCGGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 433
QY 118 GlnGlnSerAsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 137
DB 434 AGCGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 493
QY 138 AsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 157
DB 494 GAGCGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 553
QY 158 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 177
DB 554 GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 613
QY 178 ArgAlaIysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
DB 614 GAA--AGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 670
QY 197 IysGlnIleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 215
DB 671 AAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 730
QY 216 IysAsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 225
DB 731 AAGGAGCGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 760

RESULT 12
A2337339 650 bp DNA linear GSS 29-SEP-2000
LOCUS A2337339/c
DEFINITION IM0068B03F Mouse 10kb plasmid UUCGIM library Mus musculus genomic
clone UUCGIM0068B03 F, genomic survey sequence.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:28:45 ; Search time 7029.36 Seconds
(without alignments)
2555.359 Million cell updates/sec

Title: US-09-837-344-31

Perfect score: 1563
Sequence: 1 SDBEQERRAKEKIQEQOSD.....GHLEKKDGSIKPEQKEDKS 316

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O/cgn2.1/USPTO.spool.h/US09837344/runat.05122005.133532.734/app.query.faeta_1.718
-DB=GenEmbl -QMT=fastestp -SUPFIX=rge -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -STAR=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09837344 @CGN 1.1 5142 @runat.05122005.133532.734 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBSTR -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MANY TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*

1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_pro:*
10: gb_rsv:*
11: gb_by:*
12: gb_un:*
13: gb_vl:*
14: gb_vtg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1549	99.1	950	6	A28740
2	1549	99.1	950	6	ARI62961
3	1549	99.1	950	6	ARI78250

4	1549	99.1	1482	6	ARI62970	ARI62970 Sequence
5	1549	99.1	1482	6	ARI78259	ARI78259 Sequence
6	1549	99.1	1482	6	PELSA13	230320 P.falciparu
7	1546	98.9	1482	6	ARI62974	ARI62974 Sequence
8	1546	98.9	1482	6	ARI78263	ARI78263 Sequence
9	1494	95.6	5970	2	PELSA1G	X56203 P.falciparu
10	1465	93.7	1493	6	A28743	A28743 cDNA for IS
11	1465	93.7	1493	6	ARI62967	ARI62967 Sequence
12	1465	93.7	1493	6	ARI78256	ARI78256 Sequence
13	1453.5	93.0	253001	2	AE014834	AE014834 Plasmid
14	771.5	49.4	954	6	ARI62966	ARI62966 Sequence
15	771.5	49.4	954	6	ARI78255	ARI78255 Sequence
16	771.5	49.4	988	6	PELSA15	Z30319 P.falciparu
17	771.5	49.4	988	6	A28742	A28742 cDNA for IS
18	771.5	49.4	988	6	ARI62963	ARI62963 Sequence
19	771.5	49.4	988	6	ARI78252	ARI78252 Sequence
20	740.5	47.4	1368	11	AY751501	AY751501 Synthetic
21	554	35.4	843	2	PELSA1C	L40885 Plasmidium
22	554	35.4	843	2	PELSA1H	L40890 Plasmidium
23	554	35.4	843	2	PELSA1I	L40891 Plasmidium
24	553	35.4	843	2	PELSA1B	L40884 Plasmidium
25	553	35.4	843	2	PELSA1E	L40887 Plasmidium
26	553	35.4	843	2	PELSA1F	L40888 Plasmidium
27	553	35.4	843	2	PELSA1K	L40893 Plasmidium
28	550	35.2	843	2	PELSA1G	L40889 Plasmidium
29	550	35.2	843	2	PELSA1A	L40834 Plasmidium
30	549	35.1	843	2	PELSA1B	L40909 Plasmidium
31	549	35.1	843	2	PELSA1C	L40910 Plasmidium
32	549	35.1	843	2	PELSA1N	L40947 Plasmidium
33	549	35.1	843	2	PELSA1N	L40866 Plasmidium
34	549	35.1	843	2	PELSA1D	L40892 Plasmidium
35	549	35.1	843	2	PELSA1J	L40835 Plasmidium
36	549	35.1	843	2	PELSA1M	L40836 Plasmidium
37	549	35.1	843	2	PELSA1N	L40837 Plasmidium
38	549	35.1	843	2	PELSA1O	L40838 Plasmidium
39	469	30.0	795	2	AF246996	AF246996 Plasmidium
40	423	27.1	170771	5	BX465202	BX465202 Zebrafish
41	420.5	26.9	194351	5	AL772295	AL772295 Zebrafish
42	420.5	26.9	205118	5	AL928674	AL928674 Zebrafish
43	408	26.1	214506	14	AC159753	AC159753 Bos tauru
44	407.5	26.1	220046	9	AC124587	AC124587 Mus muscu
45	397	25.4	82400	9	AC090495	AC090495 Genomic s

ALIGNMENTS

RESULT 1	A28740	950 bp	DNA	linear	PAT 04-JUN-1995
LOCUS	A28740				
DEFINITION	cDNA for LSA-R-NR protein.				
ACCESSION	A28740				
VERSION	A28740.1	GI:1247512			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 950)				
TITLE	LYSER-STAGE-SPECIFIC PEPTIDE SEQUENCES OF P. FALCIPARUM BEARING				
JOURNAL	EPITOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES				
FEATURES	Patent: WO 921384-A 29 20-AUG-1992;				
source	Location/Qualifiers				
	1..950				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN
Alignment Scores: 1.01e-85 Length: 950
Pred. No.: 1549.00 Matches: 314
Score:

Percent Similarity: 99.37%
Best Local Similarity: 99.37%
Query Match: 99.10%
DB: 6
Gaps: 0

US-09-837-344-31 (1-316) x A28740 (1-950)

```
QY 1 SerAapLeuGluGlnGluArgAlaLyseGluLyseGluGlnGlnSerAapLeu 20
DB 3 AGCGATCTTAGAACAGAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 62
QY 21 GlnGlnAapArgLeuAlaLyseGluLyseGluGlnGlnSerAapLeuGlnGln 40
DB 63 GAAACAAGATGACCTGCTTAAGAAAGTTGCAAGAGCAGCAAGCATTTAGAACAGAG 122
QY 41 ArgArgAlaLyseGluLyseGluGlnGlnSerAapLeuGlnGlnArgAla 60
DB 123 AGACTTGCTTAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGACGCTGCT 182
QY 61 LyseGluLyseGluGlnGlnSerAapLeuGlnGlnArgArgAlaLyseGluLyse 80
DB 183 AAAGAAAGTTGCAAGAACAGAGACGCTGCTTAAGAACAGAGACGCTTAAGAAAG 242
QY 81 LeuGlnGlnGlnGlnSerAapLeuGlnGlnAapArgLeuAlaLyseGluLyseGln 100
DB 243 TTGCAAGAACAAAGCGATTAGAACAGATAGACTTGCTTAAGAAAGTTAGAACAGAG 302
QY 101 GlnGlnSerAapLeuGlnGlnGlnArgArgAlaLyseGluLyseGlnGlnGlnSer 120
DB 303 CAGCAAGACGATTAGAACAGAGACGCTGCTTAAGAAAGTTGCAAGAACAAAGCG 362
QY 121 AapLeuGlnGlnGlnArgArgAlaLyseGluLyseGlnGlnGlnSerAapLeuGln 140
DB 363 GATTAGAACAGAGACGCTGCTTAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAG 422
QY 141 GlnGlnArgArgAlaLyseGluLyseGlnGlnGlnSerAapLeuGlnGlnGln 160
DB 423 CAAAGAGACGCTTAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGAG 482
QY 161 ArgAlaLyseGluLyseGlnGlnGlnSerAapLeuGlnGlnArgArgAlaLyse 180
DB 483 CCGCTTAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGACGCTGCTTAAG 542
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 543 GAAAGTTGCAAGAACAGAGACGCTTAAGAACAGAGACGCTTAAGAAAGTTG 602
QY 201 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 603 CAAAGACGACCAAGAGATTAGAACAGAGACGCTTAAGAAAGTTAGAACAGAGAG 662
QY 221 LyseGluLyseGluLyseGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
DB 663 AAAAGAGCAATGAGATATATATAGCAGAGATTATATGCTTAAAGAAATACCAAGCT 722
QY 241 TleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
DB 723 ATAGAACTTCATCAAGAAATGAGAGTGGATATTATATACCATCAATCTTCTTAACT 782
QY 261 GlnAapAapArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB 783 CAGAGCAACAGAGAGAGATAGAGATTCAGAGAGAAATATCTTAATAGAAAAAACAAT 842
QY 281 ArgGlnSerTleThraenValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 300
DB 843 AGAGATCTTATTAACAACAATGTTGAAGAGCAGAAAGGATATATCAATTAAGAGCATCTTGAA 902
QY 301 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 316
DB 903 GAAAGAAAGATGCTTCAATTAACAACAAGCAAGCAAGAAAGATTAATCT 950
```

RESULT 2
AR162961

```
LOCUS AR162961 950 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 32 from patent US 6270771.
ACCESSION AR162961
VERSION AR162961.1 GI:16233425
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 950)
AUTHORS Guerin-Marchand,C. and Druilhé,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
JOURNAL bearing epitopes capable of stimulating the T lymphocytes
FEATURES
source location/Qualifiers
1..950
/organism="Unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1,01e-85 Length: 950
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
DB: 6 Gaps: 0
US-09-837-344-31 (1-316) x AR162961 (1-950)
QY 1 SerAapLeuGluGlnGlnArgAlaLyseGluLyseGluGlnGlnSerAapLeu 20
DB 3 AGCGATCTTAGAACAGAGACGCTGCTTAAGAAAGTTGCAAGAACAAAGCGATTTA 62
QY 21 GlnGlnAapArgLeuAlaLyseGluLyseGluGlnGlnSerAapLeuGlnGln 40
DB 63 GAAACAAGATGACCTGCTTAAGAAAGTTGCAAGAGCAGCAAGCATTTAGAACAGAG 122
QY 41 ArgArgAlaLyseGluLyseGluGlnGlnSerAapLeuGlnGlnArgArgAla 60
DB 123 AGACTTGCTTAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGACGCTGCT 182
QY 61 LyseGluLyseGluGlnGlnSerAapLeuGlnGlnArgArgAlaLyseGluLyse 80
DB 183 AAAGAAAGTTGCAAGAACAGAGACGCTTAAGAACAGAGACGCTTAAGAAAGTTG 242
QY 81 LeuGlnGlnGlnGlnSerAapLeuGlnGlnAapArgLeuAlaLyseGluLyseGln 100
DB 243 TTGCAAGAACAAAGCGATTAGAACAGATAGACTTGCTTAAGAAAGTTAGAACAGAG 302
QY 101 GlnGlnSerAapLeuGlnGlnGlnArgArgAlaLyseGluLyseGlnGlnGlnSer 120
DB 303 CAGCAAGACGATTAGAACAGAGACGCTGCTTAAGAAAGTTGCAAGAACAAAGCG 362
QY 121 AapLeuGlnGlnGlnArgArgAlaLyseGluLyseGlnGlnGlnSerAapLeuGln 140
DB 363 GATTAGAACAGAGACGCTGCTTAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAG 422
QY 141 GlnGlnArgArgAlaLyseGluLyseGlnGlnGlnSerAapLeuGlnGlnGln 160
DB 423 CAAAGAGCCTGCTTAAGAAAGTTGCAAGAACAAAGCGATTTAAGAACAGAGAGAG 482
QY 161 ArgAlaLyseGluLyseGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 483 CCGCTTAAGAAAGTTGCAAGAACAGAGACGCTTAAGAAAGTTGCAAGAGACGCTGTAAG 542
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 543 GAAAGTTGCAAGAACAGAGACGCTTAAGAACAGAGACGCTTAAGAAAGTTG 602
QY 201 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 603 CAAAGACGACCAAGAGATTAGAACAGAGACGCTGATACGAAAAAATTTAGAAAGAG 662
```

QY 221 LysLysGluuHISGlyAspIleLeuAlaGluAspLeuTyrglyArgLeuGluIleProAla 240
DB 663 AAAAAGGAACTGGAGATATATGACAGAGATTTATATGCTTATGAAATCCAGCT 722
QY 241 ILeGluLeuProSerGluAsnGluArgGlyTyrtTyrlleProHISGlnSerSerLeuPro 260
DB 723 ATAAACTTCATCAGAAATGAAACGTGATATTATATACCAATCATCTCTTTACCT 782
QY 261 GlnAspAsnArgGlyAsnSerArgAspSerGlyIleSerIleIleGluLeuThrAsn 280
DB 783 CAGCAACAACAGAGGAAATAGAGATTCCAAGGAAATATCTATATATGAAAAACAAT 842
QY 281 ArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHISLysGluHISLeuGlu 300
DB 843 AGAAGATCTATTACCAAAATGTTGAGACGAAAGGATATACATTAAGACATCTTGA 902
QY 301 GlnLysLysAspGlySerIleLysProGluGlnLysGluAspLysSer 316
DB 903 GAAAGAAAGATGCTTCAATTAACCAAGACAAAGAAAGATTAATCT 950
RESULT 3
AR178250 950 bp DNA linear PAT 20-APR-2002
LOCUS Sequence 32 from patent US 6319502.
DEFINITION AR178250
ACCESSION AR178250 GI:20219388
VERSION AR178250.1 GI:20219388
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 950)
AUTHORS Guerlin-Marchand,C. and Druilhé,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
JOURNAL bearing epitopes capable of stimulating the T lymphocytes
FEATURES Patent: US 6319502-A 32 20-NOV-2001;
LOCATION/Qualifiers
SOURCE 1. 950
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores: 1.01e-85 Length: 950
Pred. No.: 1549.00 Matches: 314
Score: 99.37% Conservative: 0
Percent Similarity: 99.37% Mismatches: 2
Best Local Similarity: 99.10% Indels: 0
Query Match: 6 Gaps: 0
DB: 6
US-09-837-344-31 (1-316) x AR178250 (1-950)
QY 1 SerAspLeuGluGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeu 20
DB 3 AGCGATCTAGAAACAAGAGACGCTCTAAAGAAAAGTTGCCAAGAACAAAGCATTTA 62
QY 21 GlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlu 40
DB 63 GAACAAATAGACTGCTTGAAGAAAAGTTACAAAGACGCAAAAGCATTTAGAACAAAG 122
QY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArgArgAla 60
DB 123 AGACTTGTAAAGAAAGCTGTCMAAGAACAAAGCATCTAGAACAAAGAGACGCTCT 182
QY 61 LysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArgArgAlaLysGluLys 80
DB 183 AAAAGAAAAGTTGCCAAGAACAAAGCATTTAGAACAAAGAGACGCTTAAAGAAAAG 242
QY 81 LeuGlnGlnGlnGlnSerAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGlu 100
DB 243 TTGCAAGAACAAAGCATTTAGAACAAAGATAGACTTGAAGAAAAGTTACAAAG 302
QY 101 GlnGlnSerAspLeuGlnGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnSer 120

DB 303 CAGCAAGCATTTAGAACAAAGAGACGCTTAAAGAAAAGTTGCCAAGAACAAAGC 362
QY 121 AspLeuGlnGlnGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlu 140
DB 363 GATTTAAACAAGAGACGCTTAAAGAAAAGTTGCCAAGAACAAAGCATTTTGA 422
QY 141 GlnGluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArg 160
DB 423 CAAAGAGACTTCTAAAGAAAAGTTGCCAAGAACAAAGCATTTTGAAGAACAGAGA 482
QY 161 ArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArgArgAlaLys 180
DB 483 CGTCTTAAAGAAAAGTTGCCAAGAACAAAGCATTTTGAAGAACAGAGACGCTTAA 542
QY 181 GlnLysLeuGlnGlnGlnGlnGlnSerAspLeuGlnGlnGluArgArgAlaLysGluLysLeu 200
DB 543 GAAAGTTGCAAGAACAAAGCATTTTGAACCAAGAGACGCTTAAAGAAAAGCTG 602
QY 201 GlnGlnGlnGlnArgAspLeuGlnGlnGlnGlnArgGlyAlaAspThrLysLysAsnLeuGluArg 220
DB 603 CAAAGACGCAAGACGATTTTGAACCAAGAGACGCTGATACGAAAAAATTTAGAAAGA 662
QY 221 LysLysGluuHISGlyAspIleLeuAlaGluAspLeuTyrglyArgLeuGluIleProAla 240
DB 663 AAAAAGGAACTGGAGATATATGACAGAGATTTATATGCTGTTTGAATCCAGCT 722
QY 241 ILeGluLeuProSerGluAsnGluArgGlyTyrtTyrlleProHISGlnSerSerLeuPro 260
DB 723 ATAAACTTCATCAGAAATGAAACGTGATATTATATACCAATCATCTCTTTACCT 782
QY 261 GlnAspAsnArgGlyAsnSerArgAspSerGlyIleSerIleIleGluLeuThrAsn 280
DB 783 CAGCAACAACAGAGGAAATAGAGATTTCCAAAGAAATATCTATATGAAAAACAAT 842
QY 281 ArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHISLysGluHISLeuGlu 300
DB 843 AGAAGATCTATTACCAAAATGTTGAAGACGAAAGGATATACATTAAGACATCTTGA 902
QY 301 GlnLysLysAspGlySerIleLysProGluGlnLysGluAspLysSer 316
DB 903 GAAAGAAAGATGCTTCAATTAACCAAGACAAAGAAAGATTAATCT 950
RESULT 4
AR162970 1482 bp DNA linear PAT 17-OCT-2001
LOCUS Sequence 41 from patent US 6270771.
DEFINITION AR162970
ACCESSION AR162970
VERSION AR162970.1 GI:16233435
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1482)
AUTHORS Guerlin-Marchand,C. and Druilhé,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
JOURNAL bearing epitopes capable of stimulating the T lymphocytes
FEATURES Patent: US 6270771-A 41 07-AUG-2001;
LOCATION/Qualifiers
SOURCE 1. 1482
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores: 1.52e-85 Length: 1482
Pred. No.: 1549.00 Matches: 314
Score: 99.37% Conservative: 0
Percent Similarity: 99.37% Mismatches: 2
Best Local Similarity: 99.10% Indels: 0
Query Match: 6 Gaps: 0
DB: 6
US-09-837-344-31 (1-316) x AR162970 (1-1482)

[illegible]

REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 1482)
TITLE	Guerin-Marchand, C. and Druilhé, P.
JOURNAL	hepatic sequences specific for the hepatic stages of <i>P. falciparum</i>
FEATURES	peptide sequences capable of stimulating the T lymphocytes
Source	Patent: US 6319502-A 41 20-NOV-2001;
	Location/Qualifiers
	1..1482
ORIGIN	/organism="unknown"
	/mol_type="unassigned DNA"
Alignment Scores:	
Pred. No.:	1.52e-85
Score:	1549.00
Percent Similarity:	99.37%
Best Local Similarity:	99.37%
Query Match:	99.10%
DB:	6
	Gaps: 0
US-09-837-344-31 (1-316) x AR178259 (1-1482)	
QY	1 SerAePLeuGInGInGInuAaRgAaGAlaLySgGluLySLeuGInGInGInSerAePLeu
Db	13 AGCGATCTAGAAACAAGAGAGAGAGCTCTTAAGAAAAGTTGCCAAGAACAAAGGCGATT
QY	21 GInGInaPaRgLeuAlaLySgGluLySLeuGInGInGInSerAePLeuGInGInu
Db	73 GAACAAGATAGACTTGCTTAAGAAAAGTTACAAAGAGCAAGGCGATTAGAACAAAG
QY	41 AaRgAaGAlaLySgGluLySLeuGInGInGInSerAePLeuGInGInuAaRgAaGAla
Db	133 AGACTTGCTTAAGAAAAGTTGCCAAGAACAAAGGCGATTAGAACAAAGAGAGCTGCT
QY	61 LySgGluLySLeuGInGInGInSerAePLeuGInGInuAaRgAaGAlaLySgGluLyS
Db	193 AAAGAAAAGTTGCCAAGAACAAAGGCGATTAGAACAAAGAGAGAGCTCTTAAGAAAAG
QY	81 LeuGInGInGInGInSerAePLeuGInGInaPaRgLeuAlaLySgGluLySLeuGInGIn
Db	253 TTGCCAAGAACAAAGGCGATTAGAACAAAGATAGACTTGCTTAAGAAAAGTTACAAAG
QY	101 GInGInSerAePLeuGInGInuAaRgAaGAlaLySgGluLySLeuGInGInGInSer
Db	313 CAGCAAGAGATTTAGAACAAAGAGAGAGCTGCTTAAGAAAAGTTGCCAAGAACAAAGC
QY	121 AaPLeuGInGInGInuAaRgAaGAlaLySgGluLySLeuGInGInGInSerAePLeuGIn
Db	373 GATTTAGAACAAAGAGAGAGCTCTTAAGAAAAGTTGCCAAGAACAAAGGCGATTAGAA
QY	141 GInGInaRgAaGAlaLySgGluLySLeuGInGInGInSerAePLeuGInGInuAaRg
Db	433 CAAGAGAGACTTGCTTAAGAAAAGTTGCCAAGAACAAAGGCGATTAGAACAAAGAGAG
QY	161 AaRgAaGAlaLySgGluLySLeuGInGInGInSerAePLeuGInGInuAaRgAaGAlaLyS
Db	493 CCGTCTAAAGAAAAGTTGCCAAGAACAAAGGCGATTAGAACAAAGAGAGAGCTGCTTAA
QY	181 GInuLySLeuGInGInGInSerAePLeuGInGInuAaRgAaGAlaLySgGluLySLeu
Db	553 GAAAAGTTCAAGAACAAAGGCGATTAGAACAAAGAGAGAGCTCTTAAGAAAAGTTG
QY	201 GInGInGInGInaRgAaPLeuGInGInuAaRgGAlaAaPThLySgGluLySLeuGInuAaRg
Db	613 CAAGAGCAAGAACAAAGATTTAGAACAAAGAGGCGGATACGAAAAAATTAGAAAG
QY	221 LySgGluLySLeuGInGInaPLeuAaGInuAaPLeuTyRgIYAaRgLeuGInuIeProAla
Db	673 AAAAAGGAACATGAGATATATTAGACAGAGAGATTTATATAGTCTTATAGAAATACAGCT
QY	241 ILeGInuLeProSerGInuSngInuAaRgGlyTyRtyrIleProHISGInSerSerLeuPro
Db	733 ATAGAACTTCATCGAAAATGAACGTGGATATATTATACACATCAATCTTTTAACT

Oy	261	GlnhpapanaTGGlyAenSerArgApserIyeGtUleSeriIeIleglUlvThPsn	280
Dd	793	CAGGCACCAACAGGGAATGTAGAGATTCCAAAGCAAATTTCTAATAATGAAAAACAAT	852
Oy	261	ArgJuseRietHuThPsnVnAGlUvLyrgArpAPleIHetSlYsgLjHLeGlU	300
Dd	853	AGAGATCTATTACACAAATCTTGACGACGAAAGCATATACATTAAGAACACTTCGA	912
Oy	301	GluYvLySaepGlySerIElySeProGuInIlySGUlaelySseR	316
Dd	913	GAAGAAAGAGATGCTTCATTAACACGACAAAAGAAAGATTAATCT	960
RESULT 6			
LOCUS	PFLSA13	1496 bp DNA linear INV 05-SEP-1994	
DEFINITION	P.falciparum LSA-1 gene for liver stage antigen-1 (3').		
VERSION	Z30320		
KEYWORDS	Z30320.1 GI:510183		
SOURCE	Liver stage antigen-1; LSA-1 gene.		
ORGANISM	Plasmodium falciparum (malaria parasite P. falciparum)		
REFERENCE	Plasmodium falciparum		
AUTHORS	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. 1 (bases 1 to 1496) Fidock,D.A., Grais-Messe,H., Lepers,J., Brahimi,K., Benmohamed,L., Mellouk,S., Guerin-Marchand,C., Londono,A., Rahatimalala,L., Meis,U.F., Langley,G., Roussillon,C., Taitar,A. and Drullin,P. The Plasmodium falciparum liver stage antigen-1 is well conserved J. Immunol. (1994) In press 2 (bases 1 to 1496) Fidock,D.A., Grais-Messe,H., Lepers,J., Brahimi,K., Benmohamed,L., Mellouk,S., Guerin-Marchand,C., Londono,A., Rahatimalala,L., Meis,U.F., Langley,G., Roussillon,C., Taitar,A. and Drullin,P. Plasmodium falciparum liver stage antigen-1 is well conserved and contains potent B and T cell determinants J. Immunol. 153 (1), 190-204 (1994) 7515923 3 (bases 1 to 1496) Fidock,D.A. Direct Submision Submitted (25-FEB-1994) David A Fidock, Laboratory of bio-medical parasitology, Pasteur, Institute, 25, rue du Dr. Roux, PARIS, 75724 CEDEX 15, France		
JOURNAL	TITLE	location/Qualifiers	
PUBMED	FEATURES		
REFERENCE	source		
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			

[illegible]

Db	853	AGAGATCTTATACAAACAATGTTGAAGACAGAAAGGATATATCAATAAAGACATCTTGAA	912
Qy	301	GLuIyelysAaPGIySerIleIySerProGluGlnIyIySGIyAspIySer	316
Db	913	GAAGAAGAAGATGTTCAATATAAACCGACACAAAAAGAGATTAATCT	960
RESULT 7			
LOCUS	ARI62974	1482 bp	DNA
DEFINITION	Sequence 45 from patent US 6270771.		linear
ACCESSION	ARI62974		
VERSION	ARI62974.1	GI:16233440	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1482)		
TITLE	Guerin-Marchand C. and Druilhe P.		
JOURNAL	Pepitide sequences specific for the hepatic stages of P. falciparum		
FEATURES	bearing epitopes capable of stimulating the T lymphocytes		
SOURCE	Patent: US 6270771-A 45 07-AUG-2001;		
	Location/Qualifiers		
	1..1482		

```

/mol_type="unassigned DNA"

```

Alignment Scores:

Pred. No.:	2,31e-85	Length:	1482
Score:	1546.00	Matches:	313
Percent Similarity:	99.37%	Conservative:	1
Best Local Similarity:	99.05%	Mismatches:	2
Query Match:	98.91%	Indels:	0

US-09-837-344-31 (1-316) X AR162974 (1-1482)

QY	1	serAspLeuGluGlnGlnuArgArgAlaIyVsgIuIyVLeuGlnGlnGlnGlnInserrAspLeu	20
Db	13	AGCGATCTAGAACCAAGAGACGCTGCTAAAGAAAATTGCAMGACAAACCAACGATTTA	72
QY	21	GlnGlnAspArgLeuAlaIyVsgIuIyVLeuGlnGlnGlnInserrAspLeuGlnGln	40
Db	73	GAACAAGATAGACTTGCTTAAGAAAAGTTTACAAAGACACAGCAAGCATTTAGAACAAAG	132
QY	41	ArgArgAlaIyVsgIuIyVLeuGlnGlnGlnInserrAspLeuGlnGlnuArgArgAla	60
Db	133	AGACTTGCTAAAGAAAAGTTGCAGAGAACCAAGGCGATCTTGAMACAGAGACGCTGCT	192
QY	61	IyVsgIuIyVLeuGlnGlnGlnInserrAspLeuGlnGlnuArgArgAlaIyVsgIuIyV	80
Db	193	AAAGAAAAGTTSCAAGAACCAACAAACGATTTTAGAACAGAGAGACGCTGTTAAAGAAAAG	252
QY	81	LeuGlnGlnGlnGlnInserrAspLeuGlnGlnAspArgLeuAlaIyVsgIuIyVLeuGlnGln	100
Db	253	TTGCAGAGAACCAACAAAGCGATTTAGAACAGATAGACTTGCTTAAAGAAAAGTTTACAAAG	312
QY	101	GlnGlnSerAspLeuGlnGlnGlnuArgArgAlaIyVsgIuIyVLeuGlnGlnGlnInserr	120
Db	313	CAGCAAAAGCGATTTAGAACAGAGACGCTGCTTAAAGAAAAGTTGCAGAGAACCAACAAAG	372
QY	121	AspLeuGlnGlnGlnuArgArgAlaIyVsgIuIyVLeuGlnGlnGlnInserrAspLeuGln	140
Db	373	GATTTTAGAACCAAGAGACGCTGCTTAAAGAAAAGTTSCAAGAACCAACAAAGCGATTTAGAA	432
QY	141	GlnGlnuArgArgAlaIyVsgIuIyVLeuGlnGlnGlnInserrAspLeuGlnGlnuArg	160
Db	433	CAGAGAGACTTGCTTAAAGAAAAGTTGCAGAGAACCAACAAAGCATTTAGAACAGAGAG	492
QY	161	ArgAlaIyVsgIuIyVLeuGlnGlnGlnInserrAspLeuGlnGlnuArgArgAlaIyV	180
Db	493	CGTGCTAAAGAAAAGTTGCAGAGAACCAACAAAGCATTTAGAACAGAGAGACGCTGCTAA	552

Qy	181	Gluysysleuglnglmgnglmgnsersapleuglnglmglyrgaaglaalysgluylsleu	200
Db	553	GAAAGCTTGCAAGACACAAACGCACTTTAGAACAGAGAGACGCGCTAAAGAAAGATTG	612
Qy	201	Glnglmgnglmglnargaapleuglnglmglysaalaaaphrlyslvsasbneugluarq	220
Db	613	CAAGACAGACAAAGAGATTAGAAACAAAGGACGCTGATCGAAAAAAATTTTGAAAGA	672
Qy	221	lyslvsaglunhselgyasaplleleualngluabapleuylrgrlyargleugluilaaproba	240
Db	673	AAAAAGGAAACGTGAGATATATTAGCAGAGGATTTATATGCTCGTTTGAATAACCACT	732
Qy	241	llgeluieuprosesergluasengluarqglTYrTYrllrpeohisglnserseuPro	260
Db	733	ATGAACCTCCATCAGAAATGAAACGCGGATTTATATATACACATCAATCTCTTTACT	792
Qy	261	Glnasapasnrgglyaenserargaspsersllesgluleserlelelgulysrthraen	280
Db	793	CAGACAAACAGAGGAATAGTAGAGATTTCCACGAAATGCTTATATATGAAAAACAAT	852
Qy	281	Arggluserllyethrthrashvalgluglyargarapllehslybsglyhisleuglu	300
Db	853	AGAAATCTTATTAACAACAATGTTGAAGGACGAAGGCGATATACATTAAGGACATCTTGA	912
Qy	301	Gluylslyasapgluserllyelvsproglmglnlysgluasrplysers	316
Db	913	GAAAAAGAAAGATGGTTCATATAAACCAAGAACAAAGAAAGATTAATCT	960

LOCUS	ARI78263	1482 bp	DNA	linear	PAT 20-APR-2007
DEFINITION	Sequence	45 from patent US 6319502.			
ACCESSION	ARI78263				
VERSION	ARI78263.1	GI:20219401			

SOURCE Unknown.

Unclassified.
1 (bases 1 to 1482)
REFERENCE

AUTHORS Guerin-Marchand, C. and Drulhe, P.
TITLE Peptide sequences specific for the hepatic stages of *P. falciparum* bearing epitopes capable of stimulating the T Lymphocytes
JOURNAL Patent: US 6319502-A 45 20-NOV-2001;

```

      1. .1482
source

```

```
/mol_type="unassigned DNA"
```

Alignment Scores:

Pred. No.:	2, 316-85	length:	1
Score:	1546.00	Matches:	3
Percent Similarity:	99.31%	Conservative:	1
Best Local Similarity:	99.05%	Mismatches:	0
Query Match:	99.91%	Indels:	0
DB:	6	Gaps:	0

US-09-837-344-31 (1-316) x AR178263 (1-1482)

QY 1 SerapleuGInGInGluArGAlaLySgILySleuGInGInGInSerapleu 20
Db 13 AGCGATTAGAACAGAGAGACGTCTTAAAGAAAGTTGCCAAGACAAAGCGCATTTA 72
QY 21 GInGInaPaRgLeuAlaLySgILySleuGInGInGInSerapleuGInGIn 40
Db 73 GAACAAAGTAGCTTGCTTAAAGAAAGTTACAAGAGCAAGCGATTTAGAACAGAG 132
QY 41 ArGaRgAlaLySgILySleuGInGInGInSerapleuGInGInGInArGaRgAla 60
Db 133 AGACTTGTTAAAGAAAGTTGCCAAGACAAAGCGATCTAGAACAAAGAGACGTGCT 192
QY 61 LySgILySleuGInGInGInSerapleuGInGInGInArGAlaLySgILyS 80

Db	193	AAAGAAAAGTTGCCAAGAACAAACAAAGCATTTAGAACAAAGAGACGCTCTAAAGAAAAG	252
Qy	81	LeuGIingIingInserAspLeuGIingInaapArgLeuAlaYvGIuYbLeuGIingIu	100
Db	253	TTGCMAAGAACAAACAAAGCATTTAGAACAAAGATAGACTTGGCTTAAAGAAAAGTTTCAAGAG	312
Qy	101	GIingInserAspLeuGIingInaArgArgAlaYvGIuYbLeuGIingInGIingInser	120
Db	313	CAGCAAAACCGATTTAGAACAAAGAGAGAGCTCTAAAGAAAAGTTGCCAAGAACAAAGC	372
Qy	121	AspLeuGIingInaArgArgAlaYvGIuYbLeuGIingInGIingInserAspLeuGIu	140
Db	373	GATTTAGAACAAAGAGAGAGCTCTTAAAGAAAAGTTGCCAAGAACAAACGCAATTAGAA	432
Qy	141	GIingInaArgArgAlaYvGIuYbLeuGIingInGIingInserAspLeuGIingInaArg	160
Db	433	CAGAGAGACCTTGGCTTAAAGAAAAGTTGCCAAGAACAAACGCAATTAGAACAAAGAGAG	492
Qy	161	ArgAlaYvGIuYbLeuGIingInGIingInserAspLeuGIingInaArgArgAlaYv	180
Db	493	CGTCTTAAAGAAAAGTTTCCAAAGAACAAACAAAGCCGATTTAGAACAAAGAGAGCGTCTTAA	552
Qy	181	GIuYbLeuGIingInGIingInserAspLeuGIingInaArgArgAlaYvGIuYbLeu	200
Db	553	GAAAAAGTTGCCAAGAACAAACAAAGCCATTAGAACAAAGAGAGAGCTCTTAAAGAAAAGTTG	612
Qy	201	GIingInGIingInaArgAspLeuGIingInaArgYbAlaAepThrYbYvAsnLeuGIuArg	220
Db	613	CAGAGCAGCAAGAACAAAGATTTAGAACAAAGAGAAAGCTGATAGAAAAAATTTAGAAAAG	672
Qy	221	LYvYbGIuHIGIvAspIIELeuAlaGIuAspLeuTYRGIYvArgLeuGIuIIEProAla	240
Db	673	AAAAAGCAACATGAGATATATTACACAGAGCATTTATATGGCTTTGAATATCAAGCT	732
Qy	241	IIeGIuIuEProSerGIuAsnGIuArgYbTYRTYRIIEProHIGInserSerLeuPro	260
Db	733	ATAGAACTTCATCGAAAAATGAACGTGATATTATATACCAACATCAATCTTCTTACT	792
Qy	261	GIuAspAsnArgYbAsnSerArgAspSerYbGIuIIEserIIEIGIuYbThrAsn	280
Db	793	CAGGACAAACAGAGGGAATAGTAGAATTTCCAAAGAAATGTCTATATAGAAAAAACAAT	852
Qy	281	ArgGIuSerIIEthrThrAsnValGIuGIYvArgArgAspIIEHISLeYvGIYHISLeuGIu	300
Db	853	AGAGAACTCATTAACAACAATGTGTAGAGACCAAGAGGATATATACATTAAGAGCATCTTGA	912
Qy	301	GIuYbYvYbAspGIYSerIIEYbProGIuGIuYbYbGIuAspYbSer	316
Db	913	GAAGAAAGAAAGATGTTCAATTAACACCAAGACAAAGAAAGATTAATCT	960
RESULT 9			
PFLSALG			
LOCUS	PFLSALG	5970 bp	DNA linear INV 18-APR-2005
DEFINITION	P.falciparum USA-1 gene for liver stage antigen.		
ACCESSION	X56203 S75010 S75012 S75014 S75016		
VERSION	X56203.1 GI:9915		
KEYWORDS	liver stage antigen.		
SOURCE	Plasmodium falciparum (malaria parasite P. falciparum)		
ORGANISM	Plasmodium falciparum		
REFERENCE	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
AUTHORS	1 Zhu, J. and Hollingdale, M.R.		
TITLE	Structure of Plasmodium falciparum liver stage antigen-1		
JOURNAL	Mol. Biochem. Parasitol. 48 (2), 223-226 (1991)		
PUBMED	1840628		
REFERENCES	2 (bases 1 to 5970)		
AUTHORS	Zhu, J. and Hollingdale, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-NOV-1990) Zhu J., Hollingdale M, Biomedical Research		
COMMENT	Institute, 12111 Parklane Drive, Rockville, MD 20852, U.S.A		
FEATURES	On or before Jun 27, 2005 this sequence version replaced gi:241572, gi:241573, gi:241574, gi:241575.		
	Location/Qualifiers		

[illegible]

QY	41	Argprgmlalyvsglulvyleuenglulnglnseraspleuenglulngluarararala	60
Db	4459	AGACGCTTAAGAAGAAAGCTTCAGAACCAAAAGCGATTGAAACAAGAGAGACTTGCT	4518
QY	61	Lyvsglulvyleuenglulnglnseraspleuenglulngluarararalalyvsglulvys	80
Db	4519	AAAGAAAGCTTGCAAGAGCAACAAAGAGATTGAAACAAGAGAGCGTCTAAAGAAAG	4578
QY	81	Leuenglulnglnseraspleuenglulnapharaleualalyvsglulvyleuengluln	100
Db	4579	TTGCAGAAACAACAAGCGATTTAGAACCAAGAGACGTGCTAAAGAAAGCTTCAAGAA	4638
QY	101	Glnnglnseraspleuenglulnglnuargararalalyvsglulvyleuenglulnglnser	120
Db	4639	CAACAAGCGATTTTGAACAAGAGAGACTTGCTAATGAAGAAAGCTTGCAAGAGCAACAAGA	4698
QY	121	Aapleuenglulngluarararalalyvsglulvyleuenglulnglnseraspleuengluln	140
Db	4699	GATTTAGAACAAAGAGAGCGTCTAAAGAAAGCTTGCAAGAAACAACAAGCGATTTAGAA	4758
QY	141	Glnngluarararalalyvsglulvyleuenglulnglnseraspleuenglulngluar	160
Db	4759	CAAGGAGACGTGCTAAAGAAAGCTTGCAAGAAACAACAAGCGATTTAGAAACAAGAGAGA	4818
QY	161	Argalalyvsglulvyleuenglulnglnseraspleuenglulngluarararalalyv	180
Db	4819	CGTCTTAAGAAGAAAGCTTGCAAGAGCAACAAGCGATTTTGAACAAGAGAGCGTCTTAA	4878
QY	181	Gluylvyleuenglulnglnseraspleuenglulngluarararalalyvsglulvyleu	200
Db	4879	GAAAAAGTTGCAAGAAACAACAAGAGATTGAAACAAGAGAGACTTGCTAAAGAAAGCTTG	4938
QY	201	Glnnglnnglnuargaspleuenglulnglnuarglulvyslaasprhlyvleaspleuenglun	220
Db	4939	CAAGGACGCAAGAAAGATTGAAACAAGAAAGCGTGTATGCAAAAAAATTTAGAAAGA	4998
QY	221	Lyvylvsglulnhslyvyspilleuualagluaspleuetyrgluararaleuenglulnprola	240
Db	4999	AAAAAGAAACATGAGAGCTATTGACAAGCATTTATATGCTGCTTTGAAGAAATACAGCT	5058
QY	241	Ileglualeuprosergluasngluarglulvlytyrilleprohleglnserseleupro	260
Db	5059	ATAGAACTTCATCAGAAATGAACGCTGATTTATATACCAATCATCTCTTTACT	5118
QY	261	Glnasparanaraglulvsnserar	280
Db	5119	CAGGCAACAGAGGGAATATGAGAAATTCACAGAAATATCTATATATGAAAAAACAAAT	5178
QY	281	ArggluserilleththThrasnValglulvlararararararararararararararar	300
Db	5179	AGAGAACTTATTAACAACAATGTTGAAGAGCAAGCGATTTATCAATTAAGACATCTTGAA	5238
QY	301	Gluylvlyvsnaragluserillelybprogluenglulvsnarararararararararar	316
Db	5239	GAAAAAGAAAGATGCTTCAATTAACAACAACAAGAAAGAAAGATTAATCT	5286
RESULT 10			
LOCUS	A28743	1493 bp	DNA linear PAT 04-JUN-1995
DEFINITION	CDNA for LSA gene 3' (partial).		
VERSION	A28743		
KEYWORDS	A28743.1 GI:1247516		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homidae; Homo.		
REFERENCE	1 (bases 1 to 1493)		
AUTHORS	LAYER-STAGE-SPECIFIC PEPTIDE SEQUENCES OF P. FALCIPARUM BEARING		
TITLE	EPTIOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES		

JOURNAL	Patent:	WO 9213884-A 32 20-AUG-1992;
FEATURES	Location/Qualifiers	
Source	1..1493	
	/orgnlem="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"	
ORIGIN		
Alignment Scores:		
Pred. No.:	1,89e-80	1493
Score:	1465..00	311
Percent Similarity:	98.42%	Matches: 0
Best Local Similarity:	98.42%	Mismatches: 5
Query Match:	93.73%	Indels: 3
DB:	6	Gaps: 0
US-09-837-344-31 (1-316) X A28743 (1-1493)		
QY	1 SerAspLeuGluGlnGluArgArgAlaIysGluIysLeuGlnGlnGlnSerAspLeu	20
DB	13 AGGATCTGAACAAAGAGAGAGCTGCTTAAGAAAAGTTGCCAAGAACAAAGCGATTTA	72
QY	21 GlnGlnAspArgLeuAlaIysGluIysLeuGlnGlnGlnGlnSerAspLeuGlnGln	40
DB	73 GAACAAAGATGAGCTTGCTTAAGAAAAGTTAACAGAGCGCAAGCGATTTAGAACAGAG	132
QY	41 ArgArgAlaIysGluIysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgAla	60
DB	133 AGACTTGCTTA- GAAGAGTTGCCAAGAACAAAGCGATTTAGAACAAAGAGAGCGCT	191
QY	61 LysGluIysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnArgArgAlaIysGluIys	80
DB	192 AAAGAAAGTTGCCAAGAACAAAGAGAGCTGCTTAAGAACAGAGAGAGCTGCTTAAGAAAG	251
QY	81 LeuGlnGlnGlnGlnSerAspLeuGlnGlnAspArgLeuAlaIysGluIysLeuGlnGln	100
DB	252 TTGCCAAGAACAAAGCGATTTAACAAGATGAGCTGCTTAAGAAAAGTTACAAGAG	311
QY	101 GlnGlnSerAspLeuGlnGlnGlnGlnArgArgAlaIysGluIysLeuGlnGlnGlnSer	120
DB	312 CAGCAAGCGATTTAGAACAAAGAGAGAGCTGCTTAAGAAAAGTTGCCAAGAACAAAGC	373
QY	121 AspLeuGlnGlnGlnArgArgAlaIysGluIysLeuGlnGlnGlnGlnSerAspLeuGln	148
DB	372 GATTTGAACAAAGAGAGAGCTGCTTA- GAAGAGTTGCCAAGAACAAAGCGATTTAGAA	430
QY	141 GlnGlnArgArgAlaIysGluIysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArg	168
DB	431 CAAGAGAGAGCTGCTTAAGAAAAGTTGCCAAGAACAAAGCGATTTAGAACAAAGAGAG	499
QY	161 ArgAlaIysGluIysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnArgArgAlaIys	188
DB	491 CGTGCTTAAGAAAAGTTGCCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCTTA-	549
QY	181 GlnIysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnArgArgAlaIysGluIysLeu	200
DB	550 GAAAGAGTGCAGAACAAAGCGATTTAGAACAAAGAGAGAGCGCTTAAGAAAAGCTG	603
QY	201 GlnGlnGlnGlnArgAspLeuGlnGlnGlnGlnArgIysAlaAspThrIysAsnLeuGlnArg	220
DB	610 CAAGAGCGCAAGAGAGATTTAGAACAAAGAGAGCTGCTTAAGAAAAGTTAGAAAAGAG	666
QY	221 LysIysGlnGlnGlnGlnArgIysAlaAspLeuGlnGlnGlnGlnGlnGlnGlnArgAla	240
DB	670 AAAAGGAGCATGAGATATATTAGAGAGGATTTATATGCGCTTGAATATACAGCT	723
QY	241 IleGlnLeuProSerGlnGlnGlnGlnArgIysIleProIleProIleGlnSerSerLeuPro	260
DB	730 ATGAACTTCATCAAGAAAATCAAGCTGATTTATATACCAATCAATCTTCTTTACCT	783
QY	261 GlnAspAsnArgGlnIysAsnSerArgAspSerIysGlnIleSerIleIleGlnIysThrAsn	280
DB	790 CAGGACAAACGAGAGGATATGAGAGTTCCAAAGGAAATATCTATATATGAAAAACAAAT	845

OY 281 ArgGluser11eThrThraenValGluGlyArgArgAsp11ehi;alyeglyhi;leuGlu 300
|||||
DB 850 AGAGATCTATTACACAAATGTTGAAGAGACGAAGGATTTACATTAAGACATCTTGAA 909
OY 301 GluLysLysAspGlySer11elysProGluGlnLysGluAspLysSer 316
|||||
DB 910 GAAAGAAAGATGCTCAATTAACCAAGACAAAGAAAGATTAATCT 957

RESULT 11
AR162967 1493 bp DNA linear PAT 17-OCT-2001
AR162967
LOCUS
DEFINITION Sequence 38 from patent US 6270771.
ACCESSION AR162967
VERSION AR162967.1 GI:16233432
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1493)
AUTHORS Guerin-Marchand,C. and Druilhe,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6270771-A 38 07-AUG-2001;
FEATURES
Source Location/Qualifiers
1..1493
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.89e-80 Length: 1493
Score: 1465.00 Matches: 311
Percent Similarity: 98.42% Conservative: 0
Best Local Similarity: 98.42% Mismatches: 5
Query Match: 93.73% Indels: 3
DB: Gaps: 0

US-09-837-344-31 (1-316) x AR162967 (1-1493)

OY 1 SerAspLeuGluGlnGluArgArgAlaLysGluLysLeuGlnGlnInserAspLeu 20
|||||
DB 13 AGCATCTTAGACAAAGACGCTGCTTAAGAAAGTTGCCAAGAACAAAGCGATTTA 72
OY 21 GluGlnAspArgLeuAlaLysGluLysLeuGlnGlnInserAspLeuGlnGlu 40
|||||
DB 73 GAACAAGATGACTTGGCTTAAGAAAGTTACAAAGCGAACGATTTTGAACAAGAG 132
OY 41 ArgArgAlaLysGluLysLeuGlnGlnInserAspLeuGlnGlnGluArgAla 60
|||||
DB 133 AGACTTGCTTAA-GAAAAGTTGCCAAGAACAAAGCGATCTTAAGAACAAAGACGCT 191
OY 61 LysGluLysLeuGlnGlnGlnInserAspLeuGlnGluArgArgAlaLysGlu 80
|||||
DB 192 AAAGAAAGTTGCCAAGAACAAAGCGATTTAGAACAAAGACGCTTAAGAAAG 251
OY 81 LeuGlnGlnGlnInserAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGlu 100
|||||
DB 252 TTGCAGAACAAAGCGATTTTGAACAAGATCTTAAAGAAAGTTTCAAGAG 311
OY 101 GlnGlnInserAspLeuGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnInser 120
|||||
DB 312 CAGCAAAAGCGATTTTGAACAAGACGCTTAAAGAAAGTTGCCAAGAACAAAGC 371
OY 121 AspLeuGlnGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnInserAspLeuGln 140
|||||
DB 372 GATTTAGAACAAAGACGCTGCTTA-GAAAAGTTGCCAAGAACAAAGCGATTTTGA 430
OY 141 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnInserAspLeuGlnGlnGlu 160
|||||
DB 431 CAAGAGACGCTTGAAGAAAGTTGCCAAGAACAAAGCGATTTTGAACAAGAGAG 490
OY 161 ArgAlaLysGluLysLeuGlnGlnGlnInserAspLeuGlnGlnGluArgArgAlaLys 180

DB 491 CGTCTTAAAGAAAGTTGCCAAGAACAAAGCGATTTTGAACAAGACGCTTAA- 549
|||||
OY 181 GluLysLeuGlnGlnGlnInserAspLeuGlnGlnGluArgArgAlaLysGluLysLeu 200
|||||
DB 550 GAAAGTTGCCAAGAACAAAGCGATTTTGAACAAGACGCTTAAAGAAAGTTG 609
OY 201 GlnGlnGlnGlnArgAspLeuGlnGlnGlnGlnArgAlaAspThrLysLysAsnLeuGluArg 220
|||||
DB 610 CAAGACGACGAAAGATTTTGAACAAGACGCTGATACGAAAGAAATTTTGAAGA 669
OY 221 LysLysGluLysGlyAsp11eleuAlaGluAspLeuTyGlyArgLeuGlnIleProAla 240
|||||
DB 670 AAAAAGAACATCGAATATATTAGCAGAGGATTTATATGCTGTTAGAAATACCAGCT 729
OY 241 IleGluLeuProSerGluAsnGluArgGlyTyTyTyTyleProHisGlnSerSerLeuPro 260
|||||
DB 730 ATGAACTTCATCACAAAAATGAAACGCTGATATATATACCATCAATCTTCTTAACCT 789
OY 261 GlnAspAsnArgGlyLysAsnArgAspSerGluLysSer11eIleGluLysThrAsn 280
|||||
DB 790 CAGACAAACGAGAGGATGATAGATTCACAGAAATATCTATTAAGAAAGAAACAAAT 849
OY 281 ArgGluser11eThrThraenValGluGlyArgArgAsp11ehi;alyeglyhi;leuGlu 300
|||||
DB 850 AGAGATCTATTACACAAATGTTGAAGACGAAAGGATTAATTAAGAGACATCTTGAA 909
OY 301 GluLysLysAspGlySer11elysProGluGlnLysGluAspLysSer 316
|||||
DB 910 GAAAGAAAGATGCTCAATTAACCAAGAACAAAGAAAGATTAATCT 957

RESULT 12
AR178256 1493 bp DNA linear PAT 20-APR-2002
AR178256
LOCUS
DEFINITION Sequence 38 from patent US 6319502.
ACCESSION AR178256
VERSION AR178256.1 GI:20219394
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1493)
AUTHORS Guerin-Marchand,C. and Druilhe,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6319502-A 38 20-NOV-2001;
FEATURES
Source Location/Qualifiers
1..1493
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.89e-80 Length: 1493
Score: 1465.00 Matches: 311
Percent Similarity: 98.42% Conservative: 0
Best Local Similarity: 98.42% Mismatches: 5
Query Match: 93.73% Indels: 3
DB: Gaps: 0

US-09-837-344-31 (1-316) x AR178256 (1-1493)

OY 1 SerAspLeuGlnGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnInserAspLeu 20
|||||
DB 13 AGCATCTTAGAACAAAGACGCTTAAAGAAAGTTGCCAAGAACAAAGCGATTTA 72
OY 21 GluGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnInserAspLeuGlnGlu 40
|||||
DB 73 GAACAAGATGACTTGGCTTAAGAAAGTTACAAAGCGAACGATTTTGAACAAGAG 132
OY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnInserAspLeuGlnGlnGluArgAla 60
|||||
DB 133 AGACTTGCTTAA-GAAAAGTTGCCAAGAACAAAGCGATCTTAAGAACAAAGACGCT 191

DB 180167 GATTATGAACAAGAGACTAGCTAAGAAAAATTACAGGGGAGCATGACATTTAGAA 180226
QY 132 -----LeuGIngluGIngluInserAspLeuGIngluArg 143
DB 180227 CGAAGCAAGCATCTAAGAAAAAGTTGCAAGAACAAAGCATTTAGAAACAAGAGA 180286
QY 144 ArgAlaLysGluLysLeuGIngluGIngluInserAspLeuGIngluArgAlaLys 163
DB 180287 CTTCGTTAAAGAAAAAGTTGCAAGAACAAAGCATTTAGAAACAAGAGAGCTGCTAAA 180346
QY 164 GluLysLeuGIngluGIngluInserAspLeuGIngluArgAlaLysGluLysLeu 183
DB 180347 GAAAAGTTGCAAGAACAAAGCATTTAGAAACAAGAGAGAGCTGCTAAAAGAAAGTTG 180406
QY 184 GIngluGIngluInserAspLeuGIngluArgAlaLysGluLysLeuGIngluGIn 203
DB 180407 CAAGAACAAACAAGCATTTAGAAACAAGAGAGAGCTGCTAAAAGTTGCAAGAGCAG 180466
QY 204 GlnAspAspLeuGIngluGIngluArgAlaLysGluLysLeuGIngluArgLysGlu 223
DB 180467 CAAGAGATTTGCAAGAACAAAGCATTTAGAAACAAGAAAAATTTGAAAGAAAAAGGAA 180526
QY 224 HisGluAspLeuLysLeuGluAspLeuArgLysGluLysLeuGluLysLeu 243
DB 180527 CATGGAGATGATTTAGCAAGAGATTTATATGTCGTTTGAATATCACAGCTATAGAACTT 180586
QY 244 ProSerGluAsnGluArgGlyTyrTyrIleProHisGlnSerSerLeuProGlnAspAsn 263
DB 180587 CCATCAGAAAAATGAACGTGGAATATATATACCAACATCATCTTTTACCTGAGAGCAAC 180646
QY 264 ArgGluAsnSerArgAspSerLysGluLysSerIleIleGluLysThrAsnArgLysSer 283
DB 180647 AAGGGCAATATGATGACATTCGAAAGCAAAATATCTATATGAAAAATACAAATGAGAAAT 180706
QY 284 IleThrThrAsnValGluLysArgAspIleHisLysGluLysLeuGluLysLys 303
DB 180707 ATTACACAACAATGTTGAAAGAGCAAGAGGATATACATTAAGGACATCTTGAAGAAAA 180766
QY 304 AspGlySerIleLysProGluGlnLysGluAspLysSer 316
DB 180767 GATGTTCAATTAACCAAGAACAAAAAGAAATTAATCT 180805

RESULT 14
LOCUS ARI62966 954 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 37 from patent US 6270771.
ACCESSION ARI62966
VERSION ARI62966.1 GI:16233431
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 954)
AUTHORS Guerin-Marchand,C. and Drullhe,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6270771-A 37 07-AUG-2001;
FEATURES
source Location/Qualifiers
1..954
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores: 1.35e-38 Length: 954
Pred. No.: 771.50 Matches: 163
Score: 92.59% Conservative: 12
Percent Similarity: 86.24% Mismatches: 13
Best Local Similarity: 49.36% Indels: 1
Query Match: 6 Gaps: 1
DB: 6

US-09-837-344-31 (1-316) x ARI62966 (1-954)

QY 2 AspLeuGIngluGIngluArgAlaLysGluLysLeuGIngluGIngluInserAspLeuGlu 21
DB 391 GATGACGATTAATAAAAAATATATTAAGGCGCAAGCGAAACAAACAGCAAGATCTTGA 450
QY 22 GlnAspArgLeuAlaLysGluLysLeuGIngluGIngluInserAspLeuGIngluArg 41
DB 451 ---GAAAAGCATTTAAGAAAAAGTTACAGGGGCAACAAAGCATTTCAAGAACAGAGA 507
QY 42 ArgAlaLysGluLysLeuGIngluGIngluInserAspLeuGIngluArgAlaLys 61
DB 508 CGTGCTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGACTTCTAAA 567
QY 62 GluLysLeuGIngluGIngluInserAspLeuGIngluArgAlaLysGluLysLeu 81
DB 568 GAAAAGTTGCAAGAACAAAGCATTTAGAACAAAGAGAGAGCGTCTAAAAGAAAGTTG 627
QY 82 GIngluGIngluInserAspLeuGIngluAspArgLeuAlaLysGluLysLeuGIngluGIn 101
DB 628 CAAGAACAAACAAGCATTTAGAACAAAGAGAGAGACTTCTAAAAGAAAGTTGCAAGAACAA 687
QY 102 GlnSerAspLeuGIngluGIngluArgAlaLysGluLysLeuGIngluGIngluInserAsp 121
DB 688 CAAGCGATTTAGAACAAAGAGAGAGCGTCTAAAAGAAAGTTGCAAGAACAAAGCGAT 747
QY 122 LeuGIngluGIngluArgAlaLysGluLysLeuGIngluGIngluInserAspLeuGlu 141
DB 748 TTAGAACAAAGAGAGCGTCTTAAGAAAGTTGCAAGAACAAAGCATTTAGAACAA 807
QY 142 GluArgArgAlaLysGluLysLeuGIngluGIngluInserAspLeuGIngluArg 161
DB 808 GAGAGACTTGCTTAAGAAAAAGTTACAAAGAGAGCAAAAGCGATTTAGAACAAAGATAGACTT 867
QY 162 AlaLysGluLysLeuGIngluGIngluInserAspLeuGIngluArgAlaLysGlu 181
DB 868 GCTTAAGAAAAAGTTGCAAGAACAAAGCATTTAGAACAAAGAGAGCGTCTAAAAGAA 927
QY 182 LysLeuGIngluGIngluInserAspLeu 190
DB 928 AGCTTGCAAGAACAAAGCATTTA 954

RESULT 15
LOCUS ARI78255 954 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 37 from patent US 6319502.
ACCESSION ARI78255
VERSION ARI78255.1 GI:20219393
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 954)
AUTHORS Guerin-Marchand,C. and Drullhe,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6319502-A 37 20-NOV-2001;
FEATURES
source Location/Qualifiers
1..954
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores: 1.35e-38 Length: 954
Pred. No.: 771.50 Matches: 163
Score: 92.59% Conservative: 12
Percent Similarity: 86.24% Mismatches: 13
Best Local Similarity: 49.36% Indels: 1
Query Match: 6 Gaps: 1
DB: 6

US-09-837-344-31 (1-316) x ARI78255 (1-954)

QY 2 AspLeuGIngluGIngluArgAlaLysGluLysLeuGIngluGIngluInserAspLeuGlu 21
DB 391 GATGACGATTAATAAAAAATATATTAAGGCGCAAGCGAAACAAACAGCAAGATCTTGA 450
QY 22 GlnAspArgLeuAlaLysGluLysLeuGIngluGIngluInserAspLeuGIngluArg 41
DB 451 ---GAAAAGCATTTAAGAAAAAGTTACAGGGGCAACAAAGCATTTCAAGAACAGAGA 507
QY 42 ArgAlaLysGluLysLeuGIngluGIngluInserAspLeuGIngluArgAlaLys 61
DB 508 CGTGCTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGACTTCTAAA 567
QY 62 GluLysLeuGIngluGIngluInserAspLeuGIngluArgAlaLysGluLysLeu 81
DB 568 GAAAAGTTGCAAGAACAAAGCATTTAGAACAAAGAGAGAGCGTCTAAAAGAAAGTTG 627
QY 82 GIngluGIngluInserAspLeuGIngluAspArgLeuAlaLysGluLysLeuGIngluGIn 101
DB 628 CAAGAACAAACAAGCATTTAGAACAAAGAGAGAGACTTCTAAAAGAAAGTTGCAAGAACAA 687
QY 102 GlnSerAspLeuGIngluGIngluArgAlaLysGluLysLeuGIngluGIngluInserAsp 121
DB 688 CAAGCGATTTAGAACAAAGAGAGAGCGTCTAAAAGAAAGTTGCAAGAACAAAGCGAT 747
QY 122 LeuGIngluGIngluArgAlaLysGluLysLeuGIngluGIngluInserAspLeuGlu 141
DB 748 TTAGAACAAAGAGAGCGTCTTAAGAAAGTTGCAAGAACAAAGCATTTAGAACAA 807
QY 142 GluArgArgAlaLysGluLysLeuGIngluGIngluInserAspLeuGIngluArg 161
DB 808 GAGAGACTTGCTTAAGAAAAAGTTACAAAGAGAGCAAAAGCGATTTAGAACAAAGATAGACTT 867
QY 162 AlaLysGluLysLeuGIngluGIngluInserAspLeuGIngluArgAlaLysGlu 181
DB 868 GCTTAAGAAAAAGTTGCAAGAACAAAGCATTTAGAACAAAGAGAGCGTCTAAAAGAA 927
QY 182 LysLeuGIngluGIngluInserAspLeu 190
DB 928 AGCTTGCAAGAACAAAGCATTTA 954


```
Db      391 GATGACGATAAAAAATATATTAAAGCGACAAGAAAAACAAAGAAATCTTGAA 450
Qy      22  GlnApargleuAlalySgIuLyLeuGInGIInSerAepLeuGIuGIuArg 41
      : : : : :
Db      451 ---GAAAAAGCAGCTAAAGAAAAGTTACAGGGCGCAACAAAGCGATTCAAGAACAAAGAGAGA 507
Qy      42  ArgAlalySgIuLyLeuGInGIInSerAepLeuGIuGIuArgAlalyS 61
      : : : : :
Db      508 CGTGTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAGAGCTTGCTAAA 567
Qy      62  GluLySLeuGInGIuGIInSerAepLeuGIuGIuArgAlalySgIuLySLeu 81
      : : : : :
Db      568 GAAAAGTTGCAGAACAAACCAAGCGATTAGAACAAAGAGAGCTTGCTAAAAGAAAGTTG 627
Qy      82  GInGIuGIInSerAepLeuGIuGIuApargleuAlalySgIuLySLeuGInGIu 101
      : : : : :
Db      628 CAAGAACCAACAAAGCGATTAGAACAAAGAGAGCTTGCTAAAAGAAAGTTGCAAGAACAA 687
Qy      102 GInSerAepLeuGIuGIuArgAlalySgIuLySLeuGInGIuGIInSerAep 121
      : : : : :
Db      688 CAAAGCGATTAGAACAAAGAGAGAGCTTGCTAAAAGAAAGTTGCAAGAACAAAGCGAT 747
Qy      122 LeuGIuGIInGIuArgAlalySgIuLySLeuGInGIuGIInSerAepLeuGIu 141
      : : : : :
Db      748 TTAGAACAAAGAGAGAGCTTGCTAAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAA 807
Qy      142 GluArgAlalySgIuLySLeuGInGIuGIInSerAepLeuGIuGIuArg 161
      : : : : :
Db      808 GAGAGACTTGCTAAAGAAAAGTTACAAAGAGCAGCAAGCGATTAGAACAAAGATAGACTT 867
Qy      162 AlalySgIuLySLeuGInGIuGIInSerAepLeuGIuGIuArgAlalySgIu 181
      : : : : :
Db      868 GCTAAAGAAAAGTTGCAAGAACCAAGCGATTAGAACAAAGAGAGAGCTTGCTAAAAGAA 927
Qy      182 LySLeuGInGIuGIInSerAepLeu 190
      : : : : :
Db      928 AGGTTGCAGAACCAAGCGATTTA 954
```

Search completed: December 7, 2005, 10:33:18
Job time : 7132.36 sec8

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:22:20 ; Search time 826.854 Seconds

(without alignments)
2547.054 Million cell updates/sec

Title: US-09-837-344-31

Perfect score: 1563

Sequence: 1 SDLEQERRAKEKQEQSD.....GHLEKKDGSIKPEQKEDKS 316

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-Q/cgn2.1/USPTO.spool.h/US9837344/runat.05122005.133531.726/app.query.facta_1.718
-DB=N_GeneSeq -QFMT=fasta -SUPLEX=mg -MTMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNIT8=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR_SCORE=spc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09837344.CGN.1.1.900@runat.05122005.133531.726 -NCPU=6 -ICPU=3
-NO MMAP -LARGOUBERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_GeneSeq_21.*

1: GeneSeq1980s:*
2: GeneSeq1990s:*
3: GeneSeq2000s:*
4: GeneSeq2001s:*
5: GeneSeq2001s:*
6: GeneSeq2002s:*
7: GeneSeq2002s:*
8: GeneSeq2003s:*
9: GeneSeq2003s:*
10: GeneSeq2003s:*
11: GeneSeq2003s:*
12: GeneSeq2004s:*
13: GeneSeq2004s:*
14: GeneSeq2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1549	99.1	950	2	AAQ28115
2	1549	99.1	1496	2	AAQ28119 P.falcipa
3	942.5	60.3	1566	2	AAQ05140
4	771.5	49.4	988	2	AAQ28117

5	740.5	47.4	1371	12	ADO21941	Ado21941 LSA-NRC(H
6	740.5	38.4	1374	12	ADO21919	Ado21919 LSA-NRC(H
7	606.5	47.8	1320	2	AAQ08916	AAQ08916 Plasmodiu
8	393	25.1	30191	12	ADQ97654	Adq97654 Mouse can
9	369	23.6	3662	4	ABL05274	Ab105274 Drosophi1
10	355	22.7	3012	12	ADN12155	Adn12155 Herpesvir
11	354	22.6	110000	14	AD212814	Continuation (3 of
12	353.5	22.6	110000	14	AD212814	AD212814
13	329.5	21.1	3127	12	ADN12154	Adn12154 Kaposi's
14	329.5	21.1	4320	13	ADP12190	Adp12190 Trichhya
15	329.5	21.0	4320	14	ADM00406	Adm00406 Epithelia
16	327.5	21.0	3489	3	AAA30290	Aaa30290 Kaposi's
17	327.5	21.0	3489	6	ABA93487	Abaa93487 Nucleotid
18	327.5	21.0	3489	12	AD65095	Ad65095 HIV8 DNA
19	327.5	21.0	3489	14	AD68153	Ad68153 Kaposi's
20	327.5	21.0	3207	2	AAV73805	Aav73805 KSHV LUR
21	327.5	21.0	137507	2	AAV19941	Aav19941 KSHV long
22	327.5	21.0	137508	12	ADN12162	Adn12162 Human her
23	326	20.9	1835	2	AA91242	Aa91242 T. gondii
24	326	20.9	1835	4	AA542565	Aaa42565 T. gondii
25	326	20.9	1835	10	ADG17230	Adg17230 T. gondii
26	321	20.5	1662	4	ABL05275	Ab105275 Drosophi1
27	317.5	20.3	9551	2	AA222301	Aa222301 cDNA enco
28	315.5	20.2	4755	13	ACN37449	Actn37449 Tumour-as
29	311	19.9	2479	11	ADM03281	Adm03281 Human CDN
30	306	19.8	2517	13	ADU01785	Adu01785 Novel hum
31	306	19.6	7078	4	ABL03681	Ab103681 Drosophi1
32	306	19.6	13085	4	ABL03680	Ab103680 Drosophi1
33	303.5	19.4	26214	12	ADQ97246	Adq97246 Mouse can
34	302.5	19.4	1666	2	AAQ87587	Aaq87587 DNA encod
35	301	19.3	945	10	ADF02040	Adf02040 Bacterial
36	301	19.3	3543	10	ADF01979	Adf01979 Bacterial
37	286.5	19.0	8973	12	ADP28653	Adp28653 Human sec
38	294.5	18.8	1657	13	ADX15088	Adx15088 Plant ful
39	292	18.7	42533	8	ADA56114	Ada56114 Mouse Bat
40	292	18.7	42533	9	ADA02476	Ada02476 Mouse Bat
41	292	18.7	42533	10	ADB72215	Adb72215 Mouse Bat
42	291.5	18.7	1928	11	ADM03216	Adm03216 Human CDN
43	290.5	18.6	6617	12	ADL12539	Adl12539 Human ste
44	289	18.5	210	1	AAH80899	Aah80899 Sequence
45	289	18.5	2772	5	AA590525	Aaa90525 DNA encod

ALIGNMENTS

RESULT 1
AAQ28115
ID AAQ28115 standard; DNA; 950 BP.

XX	AC	AAQ28115;	
XX	AC	AAQ28115;	
XX	DT	25-MAR-2003 (revised)	
XX	DT	08-FEB-1993 (first entry)	
XX	DE	P.falciparum LSA-R-NR coding sequence.	
XX	KM	Malaria; hepatocyte; sporozoite; plasmodium DG 536; T-cell epitope;	
XX	KW	paludism; liver stage-specific antigen; ss.	
XX	OS	Plasmodium falciparum.	
XX	FT	Key	Location/Qualifiers
XX	FT	repeat_region	3..629
XX	FT	/*cag= a	
XX	FT	/rpt_type= TANDEM	
XX	FT	repeat_unit	3..53
XX	FT	/*cag= b	
XX	FN	W09213884-A1.	
XX	PD	20-AUG-1992.	
XX	PF	05-FEB-1992; 92MO-FR000104.	

XX 05-FEB-1991: 91FR-00001286.
XX (INSP) INST PASTEUR.
XX Guerimarchand C, Druilhe P;
XX MPI. 1992-299985/36.
XX P-PSDB; AAR26941.
XX
XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX vaccination against, treatment of and diagnosis of malaria.
XX
XX Claim 17; Fig 2; 81pp: French.
XX
XX A genomic DNA bank of P. falciparum EcoRI fragments prepared in lambda
XX gIII was used to transform E.coli. The expression library was screened
XX with human antisera against antigens of all stages of P. falciparum. The
XX library was reselected with antibodies affinity- purified on a clone
XX which was able to recognise antibodies specific to the hepatic phase.
XX About 40 clones were detected which produced a characteristic LSA
XX epitope. The clone with the largest insert (950 bases) encoded LSA-R-NR
XX containing a 12-repeat region followed by a non-repeat region. Preferred
XX antigenic polypeptides of the invention are derived from the amino acid
XX sequence of LSA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
SQ Sequence 950 BP; 456 A; 127 C; 206 G; 161 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,06e-124 Length: 950
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
DB: 2 Gaps: 0

US-09-837-344-31 (1-316) x AAQ28115 (1-950)

QY 1 SerAspLeuGluGlnGluArgAlaLysGluLysLeuGlnGlnInserAspLeu 20
DB 3 ACCGATCTAGAACAAAGAGACGCTGTAAGAAAGTTCCAAAGAACAAAGCGATTTA 62
QY 21 GlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnInserAspLeuGlnGln 40
DB 63 GAACAAAGATAGACTTCTTAAGAAAGTTTCAAGAGCAGCAAGCATTTAGAACAG 122
QY 41 ArgArgAlaLysGluLysLeuGlnGlnInserAspLeuGlnGlnGluArgAla 60
DB 123 AGACTTGTAAAGAAAGTTTGCAGAAACAACAAAGCATCTAGAACAAAGAGACGCT 182
QY 61 LysGluLysLeuGlnGlnInserAspLeuGlnGlnGluArgAlaLysGluLys 80
DB 183 AAAGAAAGTTGCAAAACAACAAAGCATTTAGAACAAAGAGACGCTTAAAGAAAG 242
QY 81 LeuGlnGluGlnInserAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGln 100
DB 243 TTGCAAGAACAAAGCGATTAGAACAAAGATGACTTCTTAAGAAAGTTTCAAGAG 302
QY 101 GlnGlnInserAspLeuGlnGlnGluArgAlaLysGluLysLeuGlnGlnInser 120
DB 303 CAGCAAGAGCATTTAGAACAAAGAGACGCTTAAAGAAAGTTTGCAGAAACAACAAAGC 362
QY 121 AspLeuGlnGlnGlnGluArgAlaLysGluLysLeuGlnGlnInserAspLeuGln 140
DB 363 GATTTAGAACAAAGAGACGCTTAAAGAAAGTTTCAAGAACAAAGCGATTAGAA 422
QY 141 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnInserAspLeuGlnGlnGlu 160
DB 423 CAAGAGAGACTTCTTAAGAAAGTTTGCAGAAACAACAAAGCATTTAGAACAGAGAA 482
QY 161 ArgAlaLysGluLysLeuGlnGlnGlnInserAspLeuGlnGlnGluArgAlaLys 180
DB 483 CGTCTTAAAGAAAGTTTGCAGAAACAACAAAGCATTTAGAACAAAGAGACGCTTAA 542

QY 181 GluLysLeuGlnGlnGlnInserAspLeuGlnGlnGluArgAlaLysGluLysLeu 200
DB 543 GAAGAGTTGCAGAAACAACAAAGCATTTAGAACAAAGAGACGCTTAAAGAAAGTTG 602
QY 201 GlnGlnGlnInserAspLeuGlnGlnGluArgAlaLysGluLysLeuGlnGln 220
DB 603 CAAGAGCAGCAAGAGATTTAGAACAAAGAGCGCTATCGAAAAAATTTAGAAAGA 662
QY 221 LysLysGluLysGluAspLeuAlaLysLeuGlnGlnGlnGluArgAlaLysLeu 240
DB 663 AAAAAGGAACATGAGATATATATAGCAGAGATTTATATGCTTTAGAAATACAGCT 722
QY 241 TleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHisGlnSerSerLeuPro 260
DB 723 ATGAGACTTCCATCGAAGAAATGAAAGTGGATTTATATATCCAGATCAATCTTCTTACCT 782
QY 261 GlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThrAsn 280
DB 783 CAGGACAAACAGAGGAGATATAGAGATTTCCAAAGAAATATCTATATAGAAACAAAT 842
QY 281 ArgGluSerIleThrThrAsnValGluLysArgAlaAspIleHisLysGluLysLeu 300
DB 843 AGAGATCTATATTAACAAGATGTTGAAGACGAAGGATATACATTAAGAGCATCTTGA 902
QY 301 GluLysLysAspGlySerIleLysProGluGlnLysGluAspLysSer 316
DB 903 GAAGAAAGATGCTTCAATTAACCAAGAACAAAGAAAGATTAATCT 950

RESULT 2

AAQ28119 standard; DNA; 1496 BP.

ID AAQ28119
XX
XX AAQ28119;
XX
XX 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX

DE P. falciparum LSA gene 3' region.

XX Malaria; hepatocyte, sporozoite; T-cell epitope; B-cell epitope;
KW paludism; liver stage-specific antigen; ss.
XX

Plasmodium falciparum.

FH Key Location/Qualifiers

FT CDS 1..1482

FT FT /*tag= C

FT FT /product= "LSA_C-terminal_region"

FT FT repeat_region 37..639

FT FT /*tag= a

FT FT repeat_unit 37..87

FT FT /*tag= b

FT FT /tpc_type= TANDEM

PN MO9213884-A1.

XX 20-AUG-1992.

XX 05-FEB-1992; 92WO-FR000104.

XX 05-FEB-1991; 91FR-00001286.

XX (INSP) INST PASTEUR.

XX Guerimarchand C, Druilhe P;

XX MPI: 1992-299985/36.

XX P-PSDB; AAR26944.

XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.
XX

PS Claim 2; Fig 8-10; 81pp; French.

PS Claim 2, Fig 8-10; 81pp; French.
XX This nucleotide sequence is the 3' part of the P. falciparum liver-stage CC specific antigen (LSA) gene. It codes for a polypeptide sequence which CC carries a T cell epitope characteristic of a protein produced in CC hepatocytes infected with P. falciparum. The polypeptide can be used in CC the preparation of vaccines against malaria. (Updated on 25-MAR-2003 CC correct PN field.)

Db	766	GAACAAGAGAGCTTGTAAAGAAAAGTTGCAGAGACAAACAAGCGATTTTGAACACAGC	825
Qy	41	ArgArgAlaLysGluLysLeuGlnGlnGlnInserAspLeuGlnGlnLysArgAla	60
Db	826	AGACGCTGTTAAAGAAAAGTTGCAGAGACAAACAAGCGATTTTGAACAAGATGACTTGC	885
Qy	61	LysGluLysLeuGlnGlnGlnInserAspLeuGlnGlnLysArgAlaLysGluLys	80
Db	886	AAAGAAAAGTTGCAGAGACAAACAAGCGATTTTGAACAAGATGACGCTTAAAGAAAAG	945
Qy	81	LeuGlnGlnGlnInserAspLeuGlnGlnAspArgLeuAlaLysGluLys-----	97
Db	946	TTGCAGAGAACAAACAAGCGATTTTGAACAAGATGACTTGTAAAGAAAAGTTGCAGTA	1005
Qy	98	-----LeuGlnGlnGlnInser	103
Db	1006	CAACAACGCAATTTGAACGAACCTAACGATCTACAGAAACGTTGCAGAGACCAACGC	1065
Qy	104	AspLeuGlnGlnGlnLysArgAlaLysGluLysLeuGlnGlnInserAspLeuGln	123
Db	1066	GATTTAGAACCAAGAGAGACGTCCTAAAGAAAAGTTGCAGAGACCAACAAGCGATTTAACA	1125
Qy	124	GlnGlnLysArgAlaLysGluLysLeuGlnGlnGlnInserAspLeuGlnGlnLys	143
Db	1126	CAAGAGACGCTTAAAGAAAAGTTGCAGAGACCAACAAGAGATTTTGAACAAGAGCA	1185
Qy	144	ArgAlaLysGluLysLeuGlnGlnGlnInserAspLeuGlnGlnLysArgAlaLys	163
Db	1186	CTTGTAAAGAAAAGTTGCAGAGACCAACAAGCGATCTTGAACAAGAGACTTGTAA	1245
Qy	164	GluLysLeuGlnGlnGlnInserAspLeuGlnGlnLysArgAlaLysGluLysLeu	183
Db	1246	CAAAAGTTGCAGAGACCAACAAGATTTTGAACAAGAGACGCTTAAAGAAAAGTTG	1305
Qy	184	GlnGlnGlnGlnInserAspLeuGlnGlnLysArgAlaLysGluLysLeuGlnGln	203
Db	1306	CAAGAACCAACAAGCGATTTTGAACAAGAGACGCTTAAAGAAAAGTTGCAGAGACAA	1365
Qy	204	GlnArgAspLeuGlnGlnGlnLys-----AlaAspThrLysAsn	217
Db	1366	CAACGCAATTTGAACCAAGAGAGACTTGTAAAGTTCGCAAGAACCAACAAGCTAT	1425
Qy	218	LeuGlnArgLysLysGlnHisGlnAspIleLeuAlaGluAspLeuTyrGlyArgLeuGln	237
Db	1426	TTAGAACCGAATTAAGCATCTACAGAAACGTTGCAAGAG-----	1464
Qy	238	IleProAlaIleGlnLeuProSerGluAsnGlnLysGlyTyrTyrIleProHisGlnSer	257
Db	1465	-----CGAACAAAGC	1473
Qy	258	SerLeuProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleGln	277
Db	1474	CATTTAAGAACCAAGAGAGACTTGTAAAGAAAAGTTGCAGAGAACCAACAAGCATTTAGA	1535
Qy	278	LysThrAsnArg 281	
Db	1534	CAAGAGAGACGT 1545	
RESULT 4			
ID	AAQ28117	standard; DNA; 988 BP.	
XX	AAQ28117;		
XX	AC		
XX	25-MAR-2003 (revised)		
XX	DT 08-FEB-1993 (first entry)		
XX	DE		
XX	P.falciparum LSA gene 5' region.		
KM	Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;		
KW	paludium; liver stage-specific antigen; ss.		
XX			

OS	Plasmodium falciparum.					
XX						
XX	Key	Location/Qualifiers				
FT	5'UTR	1..32				
FT		/tag= a				
FT	misc_feature	33..988				
FT		/tag= d				
FT	repeat_region	/note= "5' part of LSA gene"				
FT		492..988				
FT		/tag= b				
FT	repeat_unit	492..542				
FT		/tag= c				
XX						
PN	M09213884-A1.					
XX						
PD	20-AUG-1992.					
XX						
Pf	05-FEB-1992;	92MO-FR000104.				
XX						
PR	05-FEB-1991;	91FR-00001286.				
XX						
PA	(INSP) INST PASTEUR.					
PI	Guerinmarchand C, Druilhe P;					
XX						
DR	WPI, 1992-299985/36.					
XX	P-PSDB; AAR65943.					
PT	Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for					
XX	vaccination against, treatment of and diagnosis of malaria.					
PS	Claim 11; Fig 6; 81pp; French.					
CC						
XX	This nucleotide sequence is the 5' part of the P.falciparum liver-stage					
CC	specific antigen (LSA) gene. It codes for a polypeptide sequence which					
CC	carries a T cell epitope characteristic of a protein produced in					
CC	hepatocytes infected with P.falciparum. The polypeptide can be used in					
CC	the preparation of vaccines against malaria. (Updated on 25-MAR-2003 to					
XX	correct PN field.)					
SQ	Sequence 988 BP; 465 A; 120 C; 184 G; 219 T; 0 U; 0 Other;					
	Alignment Scores:					
	Pred. No.:	1.79e-57	Length:	988		
	Score:	771.50	Matches:	163		
	Percent Similarity:	92.59%	Conservative:	12		
	Best Local Similarity:	86.24%	Mismatches:	13		
	Query Match:	49.36%	Indels:	1		
	DB:	2	Gaps:	1		
US-09-837-344-31 (1-316) x AAQ28117 (1-988)						
OY	2 ASPLLEUGLNGlunlargrgla1alysglulybleunglnglninserapleuglu 21					
Dd	: : : : :					
	423 GATGACGATTAATAAAATTTTAAAGGCCAAGCAAAACAAGAACATCTTGAA 482					
OY	22 GlnAsPArGLueAlaLySGlulYbLeuGInGInGInserAsPLeuGInGuArG 41					
Dd	: : : : :					
	483 ---GA AAAAGCA GCCTTAAGAAAAGTTACAGGGGCACAAAGCGATTCCAGAA CAAGA GA 539					
OY	42 ArGa1AlYbLGluLYbLeuGInGInGInserAsPLeuGInGuArGAlaLyS 61					
Dd	: : : : :					
	540 CGTGCTAAAGAAAAGTTGCAAGAACAAACAAGCGATTTGAAACMAAGAGACTTGCTTAA 599					
OY	62 GluLYbLeuGInGInGInserAsPLeuGInGuArGArGAlaLySGluLYbLeu 81					
Dd	: : : : :					
	600 GAAAA GTTCAGAAACAACAAAGCCGATTTAGAACAAAGAGAGAGCTCTAAGAAAAGTTG 659					
OY	82 GInGInGInGInserAsPLeuGInGInAsPArGLueAlaLySGlulYbLeuGInGIn 101					
Dd	: : : : :					
	660 CAAAGAAACAACAAGCATTTTAAGACAAGAGACTTGCTTAAGAAAAGTTGCAAGAA CAA 719					
OY	102 GInserAsPLeuGInGInGuArGArGAlaLySGluLYbLeuGInGInGInserAp 121					

|||||
Db 720 CAAGGCCATTTAGAACAGAGACGCTCTAAAGAAAAGTTGCAGAGAACAAAGCGAT 779
Qy 122 LeuGIuGIuGIuAArgAlaIyGluIyLeuGIuGIuGIuGIuGIuGIuGIuGIuGIu 141
Db 780 TTGAAACAGAGAGACGCTCTAAAGAAAAGTTGCAGAGAACAAAGCGATTTAGAACAA 839
Qy 142 GIuAArgAlaIyGluIyLeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 161
Db 840 GAGGAGCTTGTCTTAAAGAAAGTTACAGAGCAGAAAGCGATTTAGAACAGATGACTT 899
Qy 162 AlAlaIyGluIyLeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 181
Db 900 GCTTAAAGAAAAGTTGCAGAGAACAAAGCGATTTAGAACAGAGACGCTCTAAAGAA 959
Qy 182 LyLeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 190
Db 960 AGCTTGCAGAGAACAAAGCGATTTA 986
RESULT 5
ADO21941
ID ADO21941 standard; DNA; 1371 BP.
XX
AC ADO21941;
XX
DT 12-AUG-2004 (first entry)
XX
DE LSA-NRC(H) construct DNA derived from Malaria parasite LSA-1.
XX
KW LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
KW parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
KW malaria parasite P. falciparum; LSA-NRC(H) construct; ds; gene.
XX
OS Plasmodium falciparum.
OS Synthetic.
XX
FH Location/Qualifiers
FT 1..1371
FT CDS /product= "LSA-NRC(H) construct protein derived from
Malaria parasite LSA-1"
XX
PN WO2004044167-A2.
XX
PD 27-MAY-2004.
XX
PF 12-NOV-2003; 2003WO-US036011.
XX
PR 12-NOV-2002; 2002US-0425719P.
XX
PA (REED-) REED ARMY INST RES WALTER.
XX
PI Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Barbara A;
XX
DR WPI; 2004-420309/39.
DR P-PSDB; ADO21942.
XX
PT Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
PS Claim 16; SEQ ID NO 25; 90pp; English.
XX
CC The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mnt construct DNA of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria

CC parasite P. falciparum LSA-1 in addition to a C-terminal His6 tag.
XX
SQ Sequence 1371 BP; 539 A; 358 C; 283 G; 191 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,18e-54 Length: 1371
Score: 740.50 Matches: 177
Percent Similarity: 67.34% Conservative: 23
Best Local Similarity: 59.60% Mismatches: 66
Query Match: 47.38% Indels: 32
DB: 12 Gaps: 7
US-09-837-344-31 (1-316) x ADO21941 (1-1371)
Qy 21 GIuGIuAspArgLeuAlaIyGluIyLeuGIuGIuGIuGIuGIuGIuGIuGIuGIu 40
Db 16 GAAAGAACGCAAAATTTATCAAAAGCAATCTCCGCGAGCTTCAAC-----AGCGGC 69
Qy 41 ArgArgAlaIyGluIyLeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 60
Db 70 AACCGATCAACAGGAAAAAGCATGAGAGAAAATGTGCTGAGCCCAACTCTAGAG 129
Qy 61 LyGluIyLeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 80
Db 130 AAGACTAAAAA-CAACGAAAAACAATAATCTTGCACAGCAAGAAAGCTGACGAT--- 185
Qy 81 LeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 100
Db 186 -----GAGCAGAGTTMAAACGTATCCAGACCACTTTMAATCCCTCT----- 230
Qy 101 GIu 120
Db 231 -----GCGCAACCTCGCGCTTCCGAGAACATCTT---TCTCAAGAAAA 272
Qy 121 AspLeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 140
Db 273 CAAGCTAACAAGAGAGCAACATGTAATCATCATCAACAGACGATGAC----- 327
Qy 140 uGIuGIuAArgAlaIyGluIyLeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 160
Db 328 -AAAAAAAATACATTTAAAGCCAGATGAAATCCGACAGAAAGACTCGAA----- 378
Qy 160 gArgAlaIyGluIyLeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 180
Db 379 -----GAAAGACCTGCTGACAGCTGACCTGACAGAGAGCCCTCGCTAA 428
Qy 180 gGIuIyLeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 200
Db 429 AGAAAGCTC-----CAGAGCGCTCGCTTAAAGAAAGCT 464
Qy 200 uGIu 220
Db 465 CAGAGACCAACAGGCGACCTGGAACAGCCAGAGCTGACACGAAAAAAACCTCGAAGC 524
Qy 220 gIyLeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 240
Db 525 CAAAAAGAACACGCGACCTGCTGAGAGACTTACGCGCGCTGGAATCCAGC 584
Qy 240 AlLeuGIu 260
Db 585 TATGAACTCCATCCGAAAGAGAACCGCGCTACTACCCACAGAGACGACCTGCC 644
Qy 260 OGInAspAspArgIyAsnSerArgAspSerLyGluIleSerIleIleGIuIyGThrAs 280
Db 645 ACAAGATTAATCGCGGAATCTCCCGCAGATGAAGAAATAGCATTCGAAAAACCAA 704
Qy 280 nArgIuSerIleThrThrAsnValGIuIyArgArgAspIleHisGIuIyHisLeuGI 300
Db 705 CCGGAAAGCACTTACCAACCAACGGAAGCGCGGACATTCACAAAGGCACTCGA 764
Qy 300 uGIuIyLeuAspGIySerIleIyPProGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 316
Db 765 AGAAAGAAAGACGCTCCATCAACAGAAAGAAAGAACAAAGC 813

[illegible][illegible]


```

XX OS Plasmodium falciparum, (pLSAREPLS.14L.1).
XX PN WO9428930-A1.
XX PD 22-DEC-1994.
XX PF 10-JUN-1994; 94WO-US006652.
XX PR 11-JUN-1993; 93US-00075783.
XX PR 09-JUN-1994; 94US-00257073.
XX PA (VIRO-) VIROGENETICS CORP.
XX PI Paolucci E, De Talsone C, Tine JA;
XX DR WPI, 1995-036113/05.
XX PT Recombinant poxvirus confg. Plasmodium DNA in non-essential region -
XX PT useful in vaccines against malaria and for prodn. of Plasmodium
XX PT Immunogens.
XX PS Claim 3; Fig 11; 183pp; English.
XX CC AA080916 is the P. falciparum LSA-1 repeatless gene cDNA sequence. New
XX CC recombinant poxviruses containing either the SERA, ABRA, P1nsp10, AMA-1,
XX CC P1825, P1816, CSP, P1fSP2, LSA-1, LSA-1 repeatless, MSA-1, MSA-1 (N-
XX CC terminal p83 or C-terminal gp42) genes, or a combination of these in non-
XX CC essential regions of their genomes are claimed. These poxviruses (pref.
XX CC with a virulence reducing genomic deletion or disruption) can be used as
XX CC vaccines against malaria and for the prodn. of Plasmodium immunogens.
XX CC These viruses provide multicomponent, multistage vaccines due to their
XX CC expression of sporozite, liver stage, blood stage and sexual stage
XX CC proteins. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-
XX CC OCT-2003 to standardise OS field)
SQ Sequence 1320 BP; 619 A; 128 C; 227 G; 346 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.88e-43 Length: 1320
Score: 606.50 Matches: 131
Percent Similarity: 63.60% Conservative: 28
Best Local Similarity: 52.40% Mismatches: 56
Query Match: 38.80% Indels: 35
DB: Gaps: 4

US-09-837-344-31 (1-316) x AA080916 (1-1320)
QY 89 G|U|C|I|N|A|P|A|T|G|C|A|C|A|A|A|T|T|C|A|A|A|G|C|T|T|T|A|G|A|A|A|T|T|G|G|T|T|C|A 108
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 88 G|A|A|A|A|A|A|T|G|A|A|A|T|C|T|A|C|T|T|G|A|G|T|G|T|C|T|T|A|A|T|-----T|C|T|A|G|G 141
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 109 A|T|G|A|T|A|T|G|T|C|A|C|A|A|A|T|T|C|A|A|A|G|C|T|T|T|A|G|A|A|A|T|T|G|G|T|T|C|A 128
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 142 A|A|T|G|A|T|A|T|G|T|C|A|C|A|A|A|T|T|C|A|A|A|G|C|T|T|T|A|G|A|A|A|T|T|G|G|T|T|C|A 201
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 129 L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U 144
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 202 A|A|A|A|A|A|A|T|A|T|A|T|A|T|A|T|A|T|T|T|G|A|T|A|G|A|T|A|A|A|G|T|T|A|A|C|A|G|T|C|T 261
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 145 -----A|L|E|Y|S 146
DB 262 A|A|T|G|T|A|A|A|A|T|G|T|C|A|C|A|A|A|T|T|C|A|A|A|G|C|T|T|T|A|G|A|A|A|T|T|G|G|T|T|C|A 321
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 147 G|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U 166
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 322 G|A|G|A|T|A|T|T|C|T|T|A|A|G|A|A|A|T|A|T|A|T|A|T|A|T|A|T|A|T|A|T|A|T|A|T|A|T|A|T|A|T|A 381
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 167 G|I|N|G|I|N|G|I|N|S|E|A|P|L|E|U|G|I|N|G|I|N|G|I|N|G|I|N|G|I|N|G|I|N|G|I|N|G|I|N|G|I 186
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 382 A|T|A|A|T|A|T|G|A|T|G|A|C|G|A|T|-----A|A|A|A|A|A|A|T|A|T|T|T|A|A|G|G|C|A|A|G|A|A|A|C|A|G|A 435
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 187 G|I|N|S|E|A|P|L|E|U|G|I|N|G|I|N|G|I|N|G|I|N|G|I|N|G|I|N|G|I|N|G|I|N|G|I|N|G|I|N|G|I 206

```

```

DB 436 C|A|G|A|A|G|A|T|T|C|T|A|A|G|A|A|A|A|G|C|C|G|C|A|T|C|T|A|A|----- 471
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 207 L|E|U|G|I|N|A|T|G|L|Y|S|A|A|P|T|I|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U 226
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 472 ---G|A|A|C|G|A|G|A|G|C|T|G|T|A|C|G|A|A|A|A|A|A|T|T|T|G|A|A|A|A|A|G|A|A|C|A|T|G|G|A|T 528
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 227 I|L|E|U|A|I|G|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U 246
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 529 G|T|A|T|T|A|G|C|A|G|A|G|A|T|T|A|T|G|T|G|T|T|G|A|A|T|A|C|C|A|G|T|A|G|A|C|T|T|C|A|T|C|A|G|A 588
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 247 A|S|N|I|U|A|G|I|Y|T|Y|T|Y|L|E|P|H|I|S|I|N|S|E|S|E|U|P|R|O|G|I|N|S|P|A|S|N|A|T|G|I|Y|A|S|N 266
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 589 A|A|T|G|A|A|C|T|G|G|A|T|T|A|T|A|T|A|T|A|C|A|C|A|T|C|A|T|C|T|T|T|A|C|T|C|A|G|A|C|A|A|C|A|G|G|A|T 648
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 267 S|E|A|T|G|A|S|P|S|E|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U|L|Y|S|E|U 286
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 649 A|G|T|A|G|A|T|T|C|C|A|G|G|A|A|T|A|T|C|T|A|T|A|T|G|A|A|A|A|A|C|A|A|T|A|G|A|A|T|C|T|A|T|T|A|C|A|A 708
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 287 A|S|N|V|A|I|G|I|Y|T|Y|T|Y|L|E|P|H|I|S|I|N|S|E|S|E|U|P|R|O|G|I|N|S|P|A|S|N|A|T|G|I|Y|A|S|N 306
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 709 A|A|T|G|T|T|G|A|G|A|G|A|G|G|A|T|A|T|A|C|A|T|A|A|A|G|A|C|A|T|C|T|G|A|A|A|A|A|G|A|A|G|A|T|G|T|T|C|A 768
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 307 I|L|E|Y|S|P|R|O|G|I|N|Y|S|E|U|L|Y|S|E|U|A|P|L|Y|S|E|S|E| 316
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 769 A|T|A|A|A|C|G|A|A|C|A|A|A|A|A|G|A|A|T|A|T|C|T 798

RESULT 8
AD097654/c
ID AD097654 standard; DNA; 30191 BP.
XX AC AD097654;
XX DT 07-OCT-2004 (first entry)
XX DE Mouse cancer associated sequence MD10-022, SEQ ID 631.
XX KM Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX OS Mus musculus.
XX PN WO2004060304-A2.
XX PD 22-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US041389.
XX PR 27-DEC-2002; 2002US-00330773.
XX PA (SAGR-) SAGRES DISCOVERY INC.
XX PI MORRIS DW, Malandro MS;
XX DR WPI; 2004-543781/52.
XX PT New isolated cancer associated nucleic acids comprising at least 10
XX PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX PT cancers such as leukemia and lymphoma.
XX PS Claim 1; SEQ ID NO 631; 199pp; English.
XX CC The present invention relates to cancer associated sequences (AD097025-
XX CC AD098004). The sequences are useful for the diagnosis, prevention and/or
XX CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX CC data for this patent did not form part of the printed specification, but
XX CC was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 30191 BP; 7666 A; 7033 C; 6989 G; 8483 T; 0 U; 20 Other;

Alignment Scores:
Pred. No.: 2.52e-23 Length: 30191
Score: 393.00 Matches: 86

```


Oy	162	AlaIyegLlybLueuGIngluGInglInserkAepLueuGIngluGIngluAArgAlaIyegLu	181
		-----GAGCAG	2004
Db	1957	--GAGCAGGAGTTAGAGGACAGCAGGAGGAGTTAGAGGAGCAG-----	
Oy	182	LybLueuGIngluGInglInserAepLueuGIngluGIngluAArgAlaIyegLlybLueuGIn	201
Db	2005	GAGTTAGAGCAGACGAGCAGGAGTTAGAGGAGCAG-----GAGCAGGAGTTAGAG	2055
Oy	202	GlucIngluAArgAepLueuGIngluGIngluAArgLysAlaAepHrLybLysAenLueuGluAArgLys	221
Db	2056	GAGCAGGAGCAGGAGTTAGAGGAGCAG-----GAGCAGGAGCAGGAGTTAGAGGAGCTG	2109
Oy	222	LybGIngluHIsGlyAepLleLueuAlaGlyAepLueuTyrgIATArgLueuGluLeProAlaIle	241
Db	2110	GAAAGACAA-----GAGCAGGAGCAGGAAAGCAGGAAATTAGAGGAGCTG	2154
Oy	242	GlueuProSerGluAenGluAArgGlyTyrrTyrlleProHIsInserSerLueuProGln	261
Db	2155	GAG-----GAGCAA	2163
Oy	262	AspAenAArgLyAenSerAArgAepSerLybGluLeSerIlelleGluLybThrAsnAArg	281
Db	2164	GAGCAGGAGCAGGAGCAGGAGCAGGAGGAGTTAGAGGAGTGAAGCAGGAAAGAG	2223
Oy	282	GlueSerlleThrAsnValGluGlyAArgAArgAepLleHIsLybGlyHIsLueuGluGln	301
Db	2224	CAGAGGTTA---GAGGAGCTGGAAGAGCAGGAAAGCAGGAGGAGTTAGAGAGCTGGAAGAG	2280
Oy	302	LybLybAepGlySerlleLybProGluGlnLybGlu	313
Db	2281	CAGGAGCAGCAGGAGGCTGGAACAGCAGGAGCAGAG	2316

	Continuation (3 of 5) of ADZ12814 from base 200001 (Murine cancer-associated genomic DNA)					
	Sequence split into 5 fragments LOCUS ADZ12814 Accession Adz12814					
WP	Fragment Name	Begin	End			
WP	ADZ12814_0	1	110000			
WP	ADZ12814_1	100001	210000			
WP	ADZ12814_2	200001	310000			
WP	ADZ12814_3	300001	410000			
WP	ADZ12814_4	400001	417026			
 Alignment Scores:						
Pred. No.:	2.25e-19	Length:	110000			
Score:	354.00	Matches:	81			
Percent Similarity:	56.73%	Conservative:	96			
Best Local Similarity:	25.96%	Mismatches:	113			
Query Match:	22.65%	Indels:	23			
DB:	14	Gaps:	4			
 US-09-837-344-31 (1-316) x ADZ12814_2 (1-110000)						
QY	4	GLUGNGIUAARGAArgAlALyBGLUylLeuGngInglInsErApLeuGlInlAsp	23			
		: : : : : : : : : : : : : : : :				
Dd	29999	GAGCGAGGAGACAGACAGACAGCAGAAAGAGCGAGGAGCGAGAGACAGAGAGAGAG	30058			
QY	24	ArGLEuAlALyBGLUylLeuGngInglInsErApLeuGlInlGuArGArla	43			
		: :				
Dd	30059	GAGGAGCGAGACACAGAGAGAGCAGAGGAGCGAGGAGCGAGAGACAGAGAGAGAG	30118			
QY	44	LyBGLUylLeuGngInglInglInsErApLeuGngInglInsArGArlALyBGLUylLys	63			
		: : : : : : : : : : : : : : : : : : : : : : : : :				
Dd	30119	CAGGAGCGAGAGAGAGAGAGCAGAGCGAGGAGCGAGGAGAGACAGAGAGAGAG	30178			
QY	64	LeuGngInglInglInsErApLeuGngInglInsArGArlALyBGLUylLeuGngInl	83			
		: : : : : : : : : : : : : : : : : : : :				
Dd	30179	GAGCGAGGAGACAGAGAGAGCAGAGCGAGGAGGAGAGACAGAGAGAGAGAGAG	30238			
QY	84	GInglInsErApLeuGngInglInsArGLyBGLUylLeuGngInglInglInsEr	103			
		: : : : : : : : : : : : : : : : : : : :				
Dd	30239	CAGGAGCGAGAGGGGAGAGAGAG-----GAGGAGCGAGGAGAGAGAGAGAGAG	30292			

[illegible]

The present sequence is the Kaposi's sarcoma-associated herpesvirus, (KSHV) Latency-associated nuclear antigen (LANA) gene. KSHV is also known as Human Herpes Virus 8 (HHV8) and belongs to the rhadino virus, or gamma -2 herpes virus class. The LANA protein is necessary for the efficient persistence of rhadino virus DNA in mammalian cells. Persistent rhadino virus infection is implicated in a variety of diseases e.g. Kaposi's Sarcoma (KS), Primary Effusion Lymphoma (PEL) and multicentric Castleman's disease. In addition, KS is a common malignancy in HIV patients. KSHV persists in host cells in a latent form. One of the few genes expressed from the latent viral DNA is LANA. LANA associates with both human chromosomes and with the rhadino virus cis-acting element (RVCAE), thereby providing a tethering function: the KSHV DNA episome is "tied" to the host chromosomes. This allows the viral DNA to persist in the host cell. The present sequence may be used to screen and identify molecules that inhibit LANA interaction with RVCAE, thereby interfering with the latency cycle of this virus. Potential antiviral treatments for the above mentioned diseases may therefore be based on LANA deregulation. (Updated on 15-SEP-2003 to standardise OS field)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:30:56 ; Search time 251.048 Seconds

(without alignments)
2237.454 Million cell updates/sec

Title: US-09-837-344-31

Perfect score: 1563
1 SDEQERRAKEKLOEQOSD.....GHLEKKDGSIKPEQKEDKS 316

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Rgapop 6.0 , Rgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cg2_1/USPTO.apool/h/US09837344/runat_05122005_133533_763/app.query.fasta_1.718
-DB=Issued_Patents_NA -ORMT=faceap -SUFFIX=ent -MINMATCH=0.1 -LOOBCI=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09837344 @CGN 1.1 237 @runat_05122005_133533_763 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBSTR -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*

- 1: /cg2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cg2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cg2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cg2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cg2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cg2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 7: /cg2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cg2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cg2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1549	99.1	950	3	US-08-098-327E-32 Sequence 32, Appl
2	1549	99.1	950	3	US-08-462-625-32 Sequence 31, Appl
3	1549	99.1	1482	3	US-08-098-327E-41 Sequence 41, Appl
4	1549	99.1	1482	3	US-08-462-625-41 Sequence 41, Appl
5	1546	98.9	1482	3	US-08-098-327E-45 Sequence 45, Appl
6	1546	98.9	1482	3	US-08-462-625-45 Sequence 45, Appl
7	1465	93.7	1493	3	US-08-098-327E-38 Sequence 38, Appl
8	1465	93.7	1493	3	US-08-462-625-38 Sequence 38, Appl
9	771.5	49.4	954	3	US-08-098-327E-37 Sequence 37, Appl

10	771.5	49.4	954	3	US-08-462-625-37 Sequence 37, Appl
11	771.5	49.4	988	3	US-08-098-327E-34 Sequence 34, Appl
12	771.5	49.4	988	3	US-08-462-625-34 Sequence 34, Appl
13	606.5	38.8	1320	2	US-08-257-073-15 Sequence 15, Appl
14	344	22.0	1827	3	US-09-270-767-1308 Sequence 1308, Ap
15	344	22.0	1827	3	US-09-270-767-16590 Sequence 16590, A
16	337.5	21.0	3489	2	US-08-728-323A-1 Sequence 1, Appl1
17	337.5	21.0	3489	2	US-09-298-568-1 Sequence 1, Appl1
18	327.5	21.0	3489	3	US-09-410-399-1 Sequence 1, Appl1
19	327.5	21.0	3489	3	US-09-834-273-1 Sequence 20, Appl
20	327.5	21.0	32207	2	US-08-770-379-20 Sequence 20, Appl
21	327.5	21.0	32207	3	US-08-757-669A-20 Sequence 20, Appl
22	327.5	21.0	32207	3	US-09-230-371A-20 Sequence 20, Appl
23	326	20.9	1835	3	US-09-216-393B-80 Sequence 80, Appl
24	319	20.4	767677	3	US-09-949-016-12317 Sequence 12147, A
25	319	20.4	767677	3	US-09-949-016-17361 Sequence 17361, A
26	317.5	20.3	9551	2	US-08-056-200-93 Sequence 93, Appl
27	317.5	20.3	9551	2	US-08-800-644-93 Sequence 2325, Ap
28	301	19.3	945	3	US-09-543-681A-2325 Sequence 2264, Ap
29	301	19.3	3543	3	US-09-543-681A-2264 Sequence 11814, A
30	296.5	19.0	29927	3	US-09-949-016-11814 Sequence 17474, A
31	296.5	19.0	29927	3	US-09-949-016-17474 Sequence 17475, A
32	296.5	19.0	29927	3	US-09-949-016-17475 Sequence 16339, A
33	296.5	19.0	37802	3	US-09-949-016-12639 Sequence 1268, App
34	290.5	18.6	6617	3	US-09-976-594-268 Sequence 1666, Ap
35	288	18.4	2093	3	US-10-104-047-1666 Sequence 928, App
36	286.5	18.3	1522	3	US-09-902-540-928 Sequence 1694, App
37	284	18.2	7463	3	US-09-902-540-1894 Sequence 11814, A
38	281.5	18.0	1965	3	US-09-949-016-11814 Sequence 17474, A
39	278.5	17.8	29927	3	US-09-949-016-17474 Sequence 17475, A
40	278.5	17.8	29927	3	US-09-949-016-17475 Sequence 12639, A
41	278.5	17.8	29927	3	US-09-949-016-17475 Sequence 55, Appl
42	278.5	17.8	37802	3	US-10-164-595-53 Sequence 897, App
43	275	17.6	4056	3	US-10-164-595-55 Sequence 897, App
44	275	17.6	4248	3	US-09-799-451-897 Sequence 897, App
45	272	17.4	6158	3	US-09-799-451-897 Sequence 897, App

ALIGNMENTS

RESULT 1
US-08-098-327E-32
Sequence 32, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESS: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300

QY	1	SerAspLeuGluGlnGlnIuaRgrAalYbGluYbLeuGlnGlnIuGlnInserAspLeu	20
Db	3	AGCGATCTAGAAcAAAGAGAGCGTGTAAAGAAAAGTTGCAGAAACAAcAAACCGATTTA	62
QY	21	GluGlnAspArgLeuAlaYbSGluYbLeuGlnGlnGlnInserAspLeuGlnGlnIu	40
Db	63	GAAcCAAGATAGACTTGCTTAAGAAAAGTTATCAAGACACGAAAGCGATTATAGAACAAAG	122
QY	41	ArgArgAlaYbSGluYbLeuGlnGlnGlnInserAspLeuGlnGlnIuaRgrAa	60
Db	123	AGACTTGCTTAAGAAAAGTTGcCAAGAAcAAcAAAGCGACTCTAGAAcAAGAGAGACGTGCT	182
QY	61	LybSGluYbLeuGlnGlnGlnInserAspLeuGlnGlnIuaRgrAalYbSGluYb	80
Db	183	AAAGAAAAGTTGCAGAAcAAcAAACCGATTATGAAcAAGAGAGACGTCTAAAGAAAAG	242
QY	81	LeuGlnGlnGlnInserAspLeuGlnGlnAspArgLeuAlaYbGluYbLeuGlnGln	100
Db	243	TTGcCAAGAAcAAcCAAGCGATTATAGAAcAAAGATAGACTTGCTTAAGAAAAGTTATCAAGAG	302
QY	101	GlnGlnInserAspLeuGlnGlnIuaRgrAaAlYbGluYbLeuGlnGlnGlnInser	120
Db	303	CAGcAAAGCGATTATAGAAcAAGAGAGCGTGTAAAGAAAAGTTGCAGAAcAAcAAAGC	362
QY	121	AspLeuGlnGlnIuaRgrAalYbSGluYbLeuGlnGlnGlnInserAspLeuGln	140
Db	363	GATTTATAGAAcAAGAGAGAGCGTCTTAAGAAAAGTTGCAGAAcAAcCAAGCGATTATAGAA	422
QY	141	GlnGlnIuaRgrAalYbSGluYbLeuGlnGlnGlnInserAspLeuGlnGlnIuaRg	160
Db	423	CAGAGAGACTTGCTTAAGAAAAGTTGcCAAGAAcAAcCAAGCGATTATAGAAcAAGAGAG	482
QY	161	ArgAlaYbSGluYbLeuGlnGlnGlnInserAspLeuGlnGlnIuaRgrAalYbS	180
Db	483	CGTGCTTAAGAAAAGTTGCAGAAcAAcCAAGCGATTATAGAAcAAGAGAGACGTGCTTAA	542
QY	181	GluYbLeuGlnGlnGlnInserAspLeuGlnGlnIuaRgrAalYbSGluYbLeu	200
Db	543	GAAAAGTTGCAGAAcAAcCAAGCGATTATAGAAcAAGAGAGAGCGTCTTAAGAAAAGTTG	602
QY	201	GlnGlnGlnGlnIuaRgrAspLeuGlnGlnIuaRgrAlaYbAspThrIYbLybAsnLeuGlnArg	220
Db	603	CAGAGACGcCAAGAGATTTAGAAcAAGAGAGAGCGCTATCGAAAAAAAATTTAGAAAAG	662
QY	221	LybLybSGluYbSGluYbAlaYbLeuAlaGluAspLeuTygIYbArgLeuGlnIleProAla	240
Db	663	AAAAAGAAcATGGAGTATATTAGcAGAGCGATTATATGGTGGTTTAAAGAAATCCAGCT	722

Score:	1.89e-154	Length:	350
Pred. No.:	1549.00	Matches:	314
Percent similarity:	99.37%	Conservative:	0

Db 193 AAAAAAAAAAGTGCAGAACAAACAAAGCGATTTAGAACAGAGAGACGTCTAAAGAAAAG 252
Qy 81 LeuGInGInGInGInSerAepLeuGInGInAepAArgLeuAlaLySGInLyLeuGInGIn 100
Db 253 TTGCAGAAACAACAAAGCGATTTAGAACAGATGACCTGCTAAAGAAAAGTTACAGAG 312
Qy 101 GInGInSerAepLeuGInGInGInAArgAlaLySGInLyLeuGInGInGInSer 120
Db 313 CAGCAAAAGCGATTTAGAACAGAGAGACGTCTAAAGAAAAGTTGCAAGAACAAAC 372
Qy 121 AepLeuGInGInGInAArgAlaLySGInLyLeuGInGInGInSerAepLeuGIn 140
Db 373 GATTTAGAACAGAGAGACGTCTAAAGAAAAGTTGCAAGAACAAACCGATTTAGAA 432
Qy 141 GInGInAArgAlaLySGInLyLeuGInGInGInSerAepLeuGInGInGIn 160
Db 433 CAAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAACCGATTTAGAACAGAGAG 492
Qy 161 AArgAlaLySGInLyLeuGInGInGInSerAepLeuGInGInGInAArgAlaLyS 180
Db 493 CGTCTAAAGAAAAGTTGCAAGAACAAACCGATTTAGAACAGAGAGACGTCTAA 552
Qy 181 GInLyLeuGInGInGInGInSerAepLeuGInGInGInAArgAlaLySGInLyLeu 200
Db 553 GAAAAGTGCAGAACAAACAAAGCGATTTAGAACAGAGAGACGTCTAAAGAAAAGTTG 612
Qy 201 GInGInGInGInAArgAepLeuGInGInAArgAlaAepThrLySGInSerLeuGIn 220
Db 613 CAAGACGACCAAAAGATTTAGAACAAAGAGCGATTAACGAAAAAATTTAGAAA 672
Qy 221 LySGInGInGInGInAepLeuAlaLySGInAepLeuGInGInAArgLeuGInLeuAla 240
Db 673 AAAAAAGAACATGAGATATATTAGCAAGAGATTTATAGTCTTTAGAAATACAGCT 732
Qy 241 ILeGInLeuProSerGInGInGInAArgGlyTyTyTyLeuProGInSerSerLeuPro 260
Db 733 ATGAGACTTCATCAAGAAAATGAAACGTGATTTATATATACCATCAATCTCTTAACT 792
Qy 261 GInAepAArgGInGInSerAArgAepSerLySGInLeuSerLeuGInGInGInThrA 280
Db 793 CAGGACAAACAGAGGAGATGAGATTCACAGAAAATATCTATATAGAAAAACAAT 852
Qy 281 AArgGInSerLeuThrAepValGInGInAArgAArgAlaLySGInLyLeuGIn 300
Db 853 AGGAACTCTATTCAACAAATGTTGAAAGCAAGAGATTTACATAAGACATCTTGA 912
Qy 301 GInLyLeuAepGInSerLeuLyProGInGInLySGInAepLySer 316
Db 913 GAAAAAGAAAGATGTTCAATTAACAGAACAAAGAAAGATTAATCT 960
RESULT 4
US-08-462-625-41
; Sequence 41, Application US/08462625
; Patent No. 6319502
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/462, 625
;/ FILING DATE: 05-JUN-1995
;/ CLASSIFICATION: 424
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 08/098, 327
;/ FILING DATE: 24-NOV-1993
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: FR 91 01286
;/ FILING DATE: 05-FEB-1991
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: McGowan, Malcolm K.
;/ REGISTRATION NUMBER: 39,300
;/ REFERENCE/DOCKET NUMBER: 010830-078
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (703) 836-6620
;/ TELEFAX: (703) 836-2021
;/ INFORMATION FOR SEQ ID NO: 41:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 1482 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: DNA (genomic)
;/ FEATURE:
;/ NAME/KEY: CDS
;/ LOCATION: 1..1482
;/ PUBLICATION INFORMATION:
;/ DOCUMENT NUMBER: WO 92/13884
;/ PUBLICATION DATE: 20-AUG-1992
;/ US-08-462-625-41
Alignment Scores:
Pred. No.: 3,366-154 Length: 1482
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
Gaps: 0
US-09-837-344-31 (1-316) x US-08-462-625-41 (1-1482)
Qy 1 SerAepLeuGInGInGInAArgAlaLySGInLyLeuGInGInGInSerAepLeu 20
Db 13 ACGGATCTTAGAACAGAGAGAGACGTCTAAAGAAAAGTTGCAAGAACAAACCGATTTA 72
Qy 21 GInGInAepAArgLeuAlaLySGInLyLeuGInGInGInSerAepLeuGInGIn 40
Db 73 GAACAAAGATGACCTGCTAAAGAAAAGTTACAAAGAGCAAGCAAGCATTTAGAACAAAG 132
Qy 41 AArgAlaLySGInLyLeuGInGInGInGInSerAepLeuGInGInGInAArgAla 60
Db 133 AACCTGCTAAAGAAAAGTTGCAAGAACAAACCGATTTAGAACAGAGAGACGTCT 192
Qy 61 LySGInLyLeuGInGInGInGInSerAepLeuGInGInGInAArgAlaLySGInLy 80
Db 193 AAGAAAAGCTTCAAGAACAAACCGATTTAGAACAGAGAGACGTCTAAAGAAAAG 252
Qy 81 LeuGInGInGInGInSerAepLeuGInGInAepAArgLeuAlaLySGInLyLeuGIn 100
Db 253 TTGCAGAAACAACAAAGCGATTTAGAACAGATGACCTGCTAAAGAAAAGTTACAGAG 312
Qy 101 GInGInSerAepLeuGInGInGInGInAArgAlaLySGInLyLeuGInGInGInSer 120
Db 313 CAGCAAAAGCGATTTAGAACAGAGAGACGTCTAAAGAAAAGTTGCAAGAACAAAC 372
Qy 121 AepLeuGInGInGInGInAArgAlaLySGInLyLeuGInGInGInSerAepLeuGIn 140
Db 373 GATTTAGAACAGAGAGACGTCTAAAGAAAAGTTGCAAGAACAAACCGATTTAGAA 432
Qy 141 GInGInAArgAlaLySGInLyLeuGInGInGInGInSerAepLeuGInGInGIn 160
Db 433 CAAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAACCGATTTAGAACAGAGAG 492

QY 161 ArgAlaIyIeGluIyIeLeuGluGluGlnIserAspLeuGluGlnIyIaIyIe 180
DB 493 CGTCTAAAGAAAAGTTGCAAGAACCAAGCATTTAGAACAAAGAGCGTCTAAA 552
QY 181 GluIyIeLeuGlnIyIeGlnIserAspLeuGluGlnIyIaIyIeGluIyIe 200
DB 553 GAAAAGTTGCAAGAACCAAGCATTTAGAACAAAGAGCGTCTAAAAGAGTTG 612
QY 201 GlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 220
DB 613 CAAGAGCAGCAAGAAAGATTTAGAACCAAGAGCGTCTAAAAGAGTTG 672
QY 221 LyIeIyIeGluIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 240
DB 673 AAAAGAGCAATGAGATATATTAGCAGAGATTTATATGCTTTAGAAATCCAGCT 732
QY 241 IlleGluIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 260
DB 733 ATAGAACTTCATCAAGAAATGAAAGTGGATATTATATCCATCAATCTTTTAACT 792
QY 261 GlnAspAsnArgIyIeAsnSerArgAspSerIyIeGlnIyIeGlnIyIeGlnIyIe 280
DB 793 CAGAGCAACAGAGGAAATAGTATGATTTCCAAAGAAATATCTATATATGAAAAACAAT 852
QY 281 ArgGluSerIleThrIleAsnValIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 300
DB 853 AGAGAACTTATTAACAACAATGTTGAAAGCAAGAGGATATATCAATAAAGCACTTTGAA 912
QY 301 GluIyIeIyIeAspGluSerIleIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 316
DB 913 GAAAAGAAAGTGTTCATTAACCAAGAACAAAGAAAGATTAATCT 960
RESULT 5
US-08-098-327E-45
Sequence 45 Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATITE STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098.327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-45
Alignment Scores:
Pred. No.: 6,97e-154 Length: 1482
Score: 1546.00 Matches: 313
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match: 98.91% Indels: 0
DB: 3 Gaps: 0
US-09-837-344-31 (1-316) x US-08-098-327E-45 (1-1482)
QY 1 SerAspLeuGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 20
DB 13 AGCATTTAGAACACAGAGACCGTCTAAAAGAAAGTTGCAAGAACCAAGAGCGATTTA 72
QY 21 GluIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 40
DB 73 GAACAAATAGACTTCTTAAAGAAAGTTTCAAGAGCAGCAAGCGATTTTAAACAAGAG 132
QY 41 ArgArgAlaIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 60
DB 133 AGACTTCTTAAGAAAGTTGCAAGAACCAAGAGCGATTTAGAACACAGAGCGTGTCT 192
QY 61 LyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 80
DB 193 AAAGAAAGTTGCAAGAACCAAGAGCGATTTTAAACAAGAGACCGTCTAAAAGAAAG 252
QY 81 LeuGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 100
DB 253 TTGCAAGAACCAAGAGCGATTTTAAACAAGAGCGATTTTAAAGAAAGTTTCAAGAG 312
QY 101 GlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 120
DB 313 CAGCAAGACGATTTTAAACAAGAGCGTCTTAAAGAAAGTTTCAAGAGCGATTTTAAAG 372
QY 121 AspLeuGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 140
DB 373 GATTTTAAACAAGAGACCGTCTTAAAGAAAGTTTCAAGAGACCAAGAGATTTTAAAG 432
QY 141 GlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 160
DB 433 CAAGAGAGACTTGTCTTAAAGAAAGTTTCAAGAACCAAGAGCGATTTTAAAGAGAGAG 492
QY 161 ArgAlaIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 180
DB 493 CGTCTAAAGAAAAGTTGCAAGAACCAAGAGCGATTTTAAAGAGAGCGTCTTAAAG 552
QY 181 GluIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 200
DB 553 GAAAAGTTGCAAGAACCAAGAGCGATTTTAAAGAACCAAGAGACCGTCTTAAAGAAAGTTG 612
QY 201 GlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 220
DB 613 CAAGAGCAGCAAGAAAGATTTAGAACCAAGAGCGTCTAAAAGAGAGTTTGAAGAGA 672
QY 221 LyIeIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 240
DB 673 AAAAGAGCAATGAGATATATTAGCAGAGATTTATATGCTTTAGAAATCCAGCT 732
QY 241 IlleGluIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIeGlnIyIe 260
DB 733 ATAGAACTTCATCAAGAAATGAAAGTGGATATTATATCCATCAATCTTTTAACT 792

QY 261 GlnaaparnarglyAasnserargapserlysgulleserillegluluystrhen 280
DB 793 CAGGCAACAGAGGAAATAGTAGAATTCACAGAAATGCTATATATGAAAAACAAT 852
QY 281 ArggluserilleThThraenValgluglyArGaAsp1leHislyesglyHisleuglu 300
DB 853 AGAGATCTATTATCAACAAATGTTGAGAGACGAGGATATACATTAAGACATCTTGAA 912
QY 301 GlulyslyeAspGlySerilleysProgluglulysgluAspLysSer 316
DB 913 GAAAGAAAGATGCTCATTAATAAACAGAAACAAAAAGATAATCT 960
RESULT 6
US-08-462-625-45
; Sequence 45, Application US/08462625
; Patent No. 6319502
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHANT, Claudine
; APPLICANT: DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,625
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/098,327
; FILING DATE: 24-NOV-1993
; PRIOR APPLICATION DATA: FR 91 01286
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1482 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1482
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; US-08-462-625-45
Alignment Scores:
Pred. No.: 6.97e-154 Length: 1482
Score: 1546.00 Matches: 313
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2

Query Match: 98.91% Indels: 0
DB: 3 Gaps: 0
US-09-837-344-31 (1-316) x US-08-462-625-45 (1-1482)
QY 1 SerAspleuglulglnuArgArGAlalySglulysleuglnglulglnSerAspleu 20
DB 13 AGCGATCTTAGAACAAGAGAGAGCTGCTTAAGAAAAAGTTGCAAGAACCAAAAGCATTTA 72
QY 21 GlulgnAaPaTgLeuAlalySglulysleuglnglulglnSerAspleuglnglu 40
DB 73 GAAACAAGTAACTTGTCTAAAGAAAAGTTACAAAGACGAAAGCATTTAGAACAGAG 132
QY 41 ArgArGAlalySglulysleuglnglulglnSerAspleuglnglulglnuArgArGAla 60
DB 133 AGACTTGTCTTAAGAAAAGTTGCCAGACACAAAGCCATCTTAGAACAGAGAGAGTCT 192
QY 61 LysglulysleuglnglulglnSerAspleuglnglulglnuArgArGAlalySglulys 80
DB 193 AAAGAAAAGTTGCCAAGAACCAAAAGCATTTAGAACAAAGAGAGAGCTGCTTAAGAAAA 252
QY 81 LeuglnglulglnSerAspleuglnglulglnAaPaTgLeuAlalySglulysleuglnglu 100
DB 253 TTGCAAGAACCAAAAGCATTTAGAACAGATGACTTGTCTTAAGAAAAAGTTACAGAG 312
QY 101 GlnGlnSerAspleuglnglulglnuArgArGAlalySglulysleuglnglulglnSer 120
DB 313 CAGCAAAAGCATTTAGAACAGAGAGAGCTGCTTAAGAAAAGTTGCCAAGAACCAAAAGC 372
QY 121 AspleuglnglulglnuArgArGAlalySglulysleuglnglulglnSerAspleuglu 140
DB 373 GATTTAGAACAAAGAGAGAGCTGCTTAAGAAAAGTTGCAAGAACCAAAAGCATTTAGAA 432
QY 141 GlnGlnuArgArGAlalySglulysleuglnglulglnSerAspleuglnglulglnuArg 160
DB 433 CAAGGAGACTTGTCTTAAGAAAAGTTGCAAGAACCAAAAGCATTTAGAACAGAGAG 492
QY 161 ArgAlalySglulysleuglnglulglnSerAspleuglnglulglnuArgArGAlalyS 180
DB 493 CGTGCTTAAGAAAAGTTGCCAGACACAAAGCCATTTAGAACAGAGAGAGCTGCTAA 552
QY 181 GlulysleuglnglulglnSerAspleuglnglulglnuArgArGAlalySglulysleu 200
DB 553 GAAAAGTTGCCAAGAACCAAAAGCATTTAGAACAAAGAGAGAGCTGCTTAAGAAAAAGTTG 612
QY 201 GlnGlnGlnGlnuArgAspleuglnglulglnuArgysAlaAaPThlylyAsnleugluArg 220
DB 613 CAAGAGCAGCAAAAGATTTAGAACAAAGAGAGCTGATACGAAAAAAATTTAGAAAGA 672
QY 221 LyslySglulysgllyAsp1leuAlaGluAaPLeuTYrGlyArgleuglu1leProAla 240
DB 673 AAAAGGAACATGAGATATATTACAGAGATTTATATGTCGTTAGAAATACAGAGT 732
QY 241 lilegluLeuProSerGluAaGlnuArgglyTYrTYrilleProHicGlnSerleuPro 260
DB 733 ATAGAACTTCATCAGAAAAATGACGTGGATTTATATACACATCAATCTTTACCT 792
QY 261 GlnaaparnarglyAasnserargapserlysgulleserillegluluystrhen 280
DB 793 CAGGCAACAGAGGAAATAGTAGAATTCACAGAAATGCTATATATGAAAAACAAT 852
QY 281 ArggluserilleThThraenValgluglyArGaAsp1leHislyesglyHisleuglu 300
DB 853 AGAGATCTATTATCAACAAATGTTGAGAGACGAGGATATACATTAAGACATCTTGAA 912
QY 301 GlulyslyeAspGlySerilleysProgluglulysgluAspLysSer 316
DB 913 GAAAGAAAGATGCTCATTAATAAACAGAAACAAAAAGATAATCT 960
RESULT 7
US-08-098-327E-38
; Sequence 38, Application US/08098327E
; Patent No. 6270771

GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burne, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098.327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-38

Alignment Scores:
Pred. No.: 2,47e-145 Length: 1493
Score: 1465.00 Matches: 311
Percent Similarity: 98.42% Conservative: 0
Beat Local Similarity: 98.42% Mismatches: 5
Query Match: 93.73% Indels: 3
DB: 3 Gaps: 0

US-09-837-344-31 (1-316) x US-08-098-327E-38 (1-1493)
QY 1 SerAspLeuGluGlnGluArgAlaLysGluLysLeuGlnGlnGlnSerAspLeu 20
DB 13 AGCGATCTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 72
QY 21 GlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGln 40
DB 73 GAACAAAGTAACTGCTTAAAGAAAGTAAAGAGAGAGAGAGAGAGAGAGAG 132
QY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnGlnGlnSerAspLeuGlnGlnGln 60
DB 133 AGACTGTCTAA-GAAAAGTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 191
QY 61 LysGluLysLeuGlnGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnGln 80
DB 192 AAAGAAAGTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 251
QY 81 LeuGlnGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnGlnGln 100

DB 252 TTGCAAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 311
QY 101 GlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
DB 312 CACCAAG 371
QY 121 AspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140
DB 372 GATTTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 430
QY 141 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnGlnGlnGlnGlnGln 160
DB 431 CAAG 490
QY 161 ArgAlaLysGluLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 491 CGTGTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 549
QY 181 GlnLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 550 GAAAAGTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 609
QY 201 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 610 CAAG 669
QY 221 LysLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
DB 670 AAAAG 729
QY 241 LLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
DB 730 ATGAACTTCCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 789
QY 261 GlnAspAsnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB 790 CAG 849
QY 281 ArgGlnSerLysLysLysLysLysLysLysLysLysLysLysLysLysLys 300
DB 850 AGAAGATCTTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 909
QY 301 GlnLysLysAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 316
DB 910 GAAAAG 957

RESULT 8
US-08-462-625-38
Sequence 38, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burne, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462.625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424

THIS PAGE BLANK (USP19)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 10:33:33 ; Search time 1107.09 Seconds
(without alignments)
2360.343 Million cell updates/sec

Title: US-09-837-344-31

Perfect score: 1563
Sequence: 1 SDEGERRAKEKUEQOSD.....GHLEKKGSGIKPEQKEDKS 316

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O/cgn2_1/USPTO.apool h/US09837344/runat 05122005 133535 839/app query.fasta_1.718
-DB=Published Applications NA Main -OPMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=10count2
-TRANS=numan40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFILE=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09837344 @CGN 1 1 1364 @runat 05122005 133535 839
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DISPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=35 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubna/US10F_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1549	99.1	950	3	US-09-837-344-32
2	1549	99.1	950	3	US-09-900-963-32
3	1549	99.1	1482	3	US-09-837-344-41
4	1549	99.1	1482	3	US-09-900-963-41
5	1546	98.9	1482	3	US-09-837-344-45
6	1546	98.9	1482	3	US-09-900-963-45
7	1465	93.7	1493	3	US-09-837-344-38
8	1465	93.7	1493	3	US-09-900-963-38

9	771.5	49.4	954	3	US-09-837-344-37	Sequence 37, Appl
10	771.5	49.4	954	3	US-09-900-963-37	Sequence 37, Appl
11	771.5	49.4	988	3	US-09-837-344-34	Sequence 34, Appl
12	771.5	49.4	988	3	US-09-900-963-34	Sequence 34, Appl
13	369	23.6	3662	10	US-11-097-143-5152	Sequence 5152, Ap
14	329.5	21.1	4320	8	US-10-754-079-159	Sequence 159, App
15	327.5	21.0	3489	3	US-09-894-773-1	Sequence 1, Appl1
16	327.5	21.0	3489	6	US-10-294-804-1	Sequence 1, Appl1
17	327.5	21.0	3489	8	US-10-194-046-1	Sequence 1, Appl1
18	326	20.9	1835	3	US-09-216-393-80	Sequence 80, Appl
19	326	20.9	1835	6	US-10-321-856-80	Sequence 80, Appl
20	321	20.5	1662	10	US-11-097-143-5153	Sequence 5153, Ap
21	311	19.9	2479	6	US-10-108-260A-1966	Sequence 1966, Ap
22	307	19.6	720	5	US-10-101-487-76	Sequence 76, Appl
23	307	19.6	720	5	US-10-101-487-76	Sequence 76, Appl
24	307	19.6	720	5	US-10-939-988-74	Sequence 74, Appl
25	307	19.6	720	9	US-10-939-988-76	Sequence 76, Appl
26	306	19.6	7078	10	US-11-097-143-2762	Sequence 2762, Ap
27	306	19.6	13085	10	US-11-097-143-2761	Sequence 2761, Ap
28	294.5	18.8	1657	7	US-10-425-114-9663	Sequence 9663, Ap
29	294.5	18.8	2406	7	US-10-424-599-53216	Sequence 53216, A
30	292	18.7	42533	6	US-10-004-113-43	Sequence 43, Appl
31	291.5	18.7	1928	6	US-10-108-260A-1901	Sequence 1901, Ap
32	289	18.5	2772	9	US-10-450-763-8591	Sequence 8591, Ap
33	289	18.5	2772	9	US-10-450-763-26329	Sequence 26329, A
34	288	18.4	1824	9	US-10-450-763-17292	Sequence 17292, A
35	288	18.4	2093	6	US-10-104-047-1666	Sequence 1666, Ap
36	288	18.4	2850	9	US-10-450-763-15499	Sequence 15499, A
37	287	18.4	2243	7	US-10-276-774-433	Sequence 433, App
38	286.5	18.3	1522	6	US-10-104-047-1093	Sequence 1093, Ap
39	286.5	18.3	42999	3	US-09-799-462A-17	Sequence 17, Appl
40	286.5	18.3	42999	3	US-09-738-610A-17	Sequence 17, Appl
41	286.5	18.3	42999	3	US-09-738-610A-17	Sequence 73, Appl
42	286.5	18.3	42999	5	US-10-125-767-17	Sequence 17, Appl
43	286.5	18.3	42999	5	US-10-151-081-17	Sequence 17, Appl
44	286.5	18.3	42999	5	US-10-287-313-17	Sequence 17, Appl
45	286.5	18.3	42999	5	US-10-219-694-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-837-344-32
Sequence 32, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUIHE, Pierre
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-837-344-32
Alignment Scores:
Pred. No.: 5,576-130 Length: 950
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
Gaps: 0
DB: 3
US-09-837-344-31 (1-316) x US-09-837-344-32 (1-950)
QY 1 SerAspLeuGluGlnGluArgArgAlaLysGluLysLeuGlnGlnInserAspLeu 20
DB 3 AGCGATCTGAACAAGAGAGAGCTGCTAAAGAAATTCGACAGAACAAAGCGATTGA 62
QY 21 GluGlnAspArgLeuAlaLysGluLysLeuGlnGlnInserAspLeuGlnGlu 40
DB 63 GAACAAGATGAGACTTCTTAAGAAAGTTACAGACAGCAAGCGATTAGAACAAAG 122
QY 41 ArgArgAlaLysGluLysLeuGlnGlnInserAspLeuGlnGlnGluArgAla 60
DB 123 AGCTTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTTAAGAAAGAGAGAGCTGCT 182
QY 61 LysGluLysLeuGlnGlnInserAspLeuGlnGlnGluArgAlaLysGluLys 80
DB 183 AAAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAGAGAGCTGCTAAAGAAAG 242
QY 81 LeuGlnGlnGlnInserAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGlu 100
DB 243 TTGCAAGAACAAAGCGATTGAAACAAAGATGAGACTTCTAAAGAAAGTTACAAAG 302
QY 101 GlnGlnInserAspLeuGlnGlnGluArgAlaLysGluLysLeuGlnGlnInser 120
DB 303 CAGCAAGCGATTGAAACAAAGAGAGCGCTAAAGAAAGTTGCAAGAACAAAGCG 362
QY 121 AspLeuGlnGlnGlnGluArgArgAlaLysGluLysLeuGlnGlnInserAspLeuGlu 140
DB 363 GATTGAAACAGAGAGAGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTGAAAG 422
QY 141 GlnGluArgArgAlaLysGluLysLeuGlnGlnInserAspLeuGlnGlnGlu 160
DB 423 CAGAGAGAGACTTCTTAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAAGAGAG 482
QY 161 ArgAlaLysGluLysLeuGlnGlnInserAspLeuGlnGlnGluArgAlaLys 180
DB 483 CGTGCTAAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAAGAGAGAGCTGCTAA 542
QY 181 GluLysLeuGlnGlnGlnInserAspLeuGlnGlnGluArgAlaLysGluLysLeu 200
DB 543 GAAGAAAGTTGCAAGAACAAAGCGATTGAAACAGAGAGAGCTGCTAAAGAAAGTTG 602
QY 201 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 603 CAGAGAGAGAGAGAGTTTGAAGAACAAAGAGAGCTGCTAAGAAAGAAATTTAGAAAG 662

QY 221 LysLysGluLysGluAspLeuAlaGluAspLeuTyrgLysLeuGlnLysProAla 240
DB 663 AAAAGAAACATCGAGATATATTAGCAGAGAGTTTATATGGTCGTTGAAATACAGCT 722
QY 241 TleGluLeuProSerGluAsnGluArgGlyTyryTyrIleProHisGlnSerSerLeuPro 260
DB 723 ATAGAACTTCATCAGAAATGAACGTGAGATATTATACCATCAATCATCTTTACT 782
QY 261 GlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleGluLysThrAsn 280
DB 783 CAGGACAAAGAGGAGATGAGATTCAGAGAAATATCTATATATGAAAAACAAAT 842
QY 281 ArgGluSerIleThrAsnValGluLysArgArgAspIleHisLysGlyHisLeuGlu 300
DB 843 AGAGAAATCTATTACAAACAAATCTTGAGAGAGCAAGAGATATCATTAAGAGCATCTTGA 902
QY 301 GluLysLysAspGlySerIleLysProGluGlnLysGluAspLysSer 316
DB 903 GAAAGAAAGATGTTCAATTAACCAAGAACAAAGAAAGATTAATCT 950
RESULT 2
US-09-900-963-32
Sequence 32, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-900-963-32
Alignment Scores:
Pred. No.: 5,576-130 Length: 950

DB 123 AGACTTGTAAAGAAAAGTTGCAAGAACAAAGCGATCTGAAACAAAGAGAGCTGCT 192
QY 61 LYSGLULYSEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 80
DB 193 AAAAGAAAAGTTCCAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCTAAAGAAAAG 252
QY 81 LEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 100
DB 253 TTGCAAGAACAAAGCGATTTAGAACAAAGTAGAGCTTGAAGAAAAGTTACAAAGG 312
QY 101 GINGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 120
DB 313 CAGCAAGGAGATTTAGAACAAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGC 372
QY 121 AEPLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 140
DB 373 GATTTAGAACAAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAA 432
QY 141 GINGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 160
DB 433 CAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAG 492
QY 161 ARGALALYSEGLULYSEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 180
DB 493 CGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCTAA 552
QY 181 GIULYSEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 200
DB 553 GAAAAGTTCCAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCTAAAGAAAAGTTG 612
QY 201 GINGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 220
DB 613 CAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAG 672
QY 221 LYSGLULYSEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 240
DB 673 AAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCTAA 732
QY 241 LLEGLULYSEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 260
DB 733 ATAGAGCTTCCATCAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAG 792
QY 261 GINAPASAPNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 280
DB 793 CAGGACAAAG 852
QY 281 ATGGGLYSEGLULYSEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 300
DB 853 AGAGAAATCTATTACAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAG 912
QY 301 GIULYSEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 316
DB 913 GAAAAGAAAAGTTGCTCAATTAAGAACAAAGAAAAGAGATTAATCT 960
RESULT 4
US-09-900-963-41
Sequence 41, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-900-963-41
Alignment Scores:
Pred. No.: 8,966-130 Length: 1482
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
Gaps: 0
US-09-837-344-31 (1-316) x US-09-900-963-41 (1-1482)
QY 1 SerAEPLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 20
DB 13 ACGGATCTTAGAACAAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 72
QY 21 GINGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 40
DB 73 GAAAGAGATAGAGCTTGAAGAAAAGTTACAAAGAGAGAGAGAGAGAGAGAGAG 132
QY 41 ARGAGALALYSEGLULYSEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 60
DB 133 AGACTTGTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCT 192
QY 61 LYSGLULYSEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 80
DB 193 AAAAGAAAAGTTCCAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCTAAAGAAAAG 252
QY 81 LEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 100
DB 253 TTGCAAGAACAAAGCGATTTAGAACAAAGTAGAGCTTGAAGAAAAGTTACAAAGG 312
QY 101 GINGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 120
DB 313 CAGCAAGGAGATTTAGAACAAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGC 372
QY 121 AEPLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 140
DB 373 GATTTAGAACAAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAA 432

Db 673 AAAAAGAAACATGAGATATATTACAGAGATTTATATGCGTTGAAATACACT 732
Qy 241 IIEGLIUEPProSeRGluBngluArglyTYTYTIIeProHieGlnSerSerleuPro 260
Db 723 ATAGAACTTCATCAGAAAATGACGTGATATATATACCAATCAATCTCTTACT 792
Qy 261 GINAAAPAAARGLYAAsnSerARgAspSerlySGIuIleSerIleIleGluYsThraSn 280
Db 793 CAGGACACAGAGGAATAGAGATTCACAGAAATGCTATATATGAAAAACAAT 852
Qy 281 ARGGLuSerIleThraSnValGluGlyARgARgAspIleHieLySGlyHieBnglu 300
Db 853 AGAGATCTATTACAAACAATGTTGAGAGCAGAGGATATACATPAAGACATCTTGA 912
Qy 301 GIuYsLyAspGlySerIleYsProGluGluYsGluAspLySer 316
Db 913 GAAAAGAAAGATGTTCAATAAAACCAAGAAAAAGAAATAAATCT 960

RESULT 6
US-09-900-963-45
Sequence 45, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUIIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Machis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 97/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-900-963-45

Alignment Scores:

	Pred. No.:	1,67e-129	Length:	1482
	Score:	1546.00	Matches:	313
	Percent Similarity:	99.37%	Conservative:	1
	Best Local Similarity:	99.05%	Mismatches:	2
	Query Match:	98.91%	Indels:	0
	DB:	3	Gaps:	0
US-09-837-344-31 (1-316) x US-09-900-963-45 (1-1482)				
Qy	1	SeRAspLeuGluGluGluARgAlaYsGlyYsLeuGluGluGlnGlnSerAspLeu	20	
Db	13	AGCGATCTTAAGAACAAAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAAGCAATTTA	72	
Qy	21	GIuGlnAspARgLeuAlaYsGlyYsLeuGluGluGlnGlnSerAspLeuGluGlu	40	
Db	73	GACAGATTAACCTTGCTAAAGAAAAGTTACAGAGCAGCAAGCAAGCAATTTAGAACAG	132	
Qy	41	ARgARgAlaYsGlyYsLeuGluGluGlnGlnSerAspLeuGluGluARgAla	60	
Db	133	AGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATCTTAAGAACAAAGAGAGCTGCT	192	
Qy	61	LYSGIuYsLeuGluGluGlnGlnSerAspLeuGluGluARgARgAlaYsGlyYs	80	
Db	193	AAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTGTAAAGAAAAG	252	
Qy	81	LeuGlnGlnGlnSerAspLeuGluGlnAspARgLeuAlaYsGlyYsLeuGluGlu	100	
Db	253	TTGCAAGAACAAAGCAATTTAGAACAGATGACTTGCTAAAGAAAAGTTACAGAG	312	
Qy	101	GlnGlnSerAspLeuGluGlnGluARgARgAlaYsGlyYsLeuGluGlnGlnSer	120	
Db	313	CAGCAAAAGCGATTTAGAACAAAGAGAGCGTGTAAAGAAAAGTTGCAAGAACAAAGC	372	
Qy	121	AspLeuGluGlnGluARgARgAlaYsGlyYsLeuGluGlnGlnSerAspLeuGlu	140	
Db	373	GATTTAGAACAAAGAGAGCGTGTAAAGAAAAGTTGCAAGAACAAAGCAATTTAGAA	432	
Qy	141	GlnGluARgARgAlaYsGlyYsLeuGluGlnGlnGlnSerAspLeuGluGluARg	160	
Db	433	CAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCAATTTAGAACAGAGAG	492	
Qy	161	ARgAlaYsGlyYsLeuGluGlnGlnGlnSerAspLeuGluGluARgARgAlaYs	180	
Db	493	CGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGTAA	552	
Qy	181	GIuYsLeuGluGlnGlnGlnSerAspLeuGluGluARgARgAlaYsGlyYsLeu	200	
Db	553	GAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTGTAAAGAAAAGTTG	612	
Qy	201	GlnGluGlnGlnARgAspLeuGluGlnARgYsAlaAspThraYsLeuBngluARg	220	
Db	613	CAGAGCAGCAAGAGATTTAGAACAAAGAGCGTGTATGCAAAAAAATTTAGAAAGA	672	
Qy	221	IYsLYsGluYsLeuGluGlnGlnGlnSerAspLeuYsGlyYsARgLeuGluIlePro	240	
Db	673	AAAAAGAAACATGAGATATATTACAGAGGATTTATATATGCTGTTGAAAATACCA	732	
Qy	241	IIEGLIUEPProSeRGluBngluArglyTYTYTIIeProHieGlnSerSerleuPro	260	
Db	733	ATAGAACTTCATCAGAAAATGACGTGATATATATACCAATCAATCTCTTACT	792	
Qy	261	GINAAAPAAARGLYAAsnSerARgAspSerlySGIuIleSerIleIleGluYsThra	280	
Db	793	CAGGACACAGAGGAATAGTAGAGTTCCAGGAAATGCTATATATGAAAAACAAT	852	
Qy	281	ARGGLuSerIleThraSnValGluGlyARgARgAspIleHieLySGlyHieBnglu	300	
Db	853	AGAGATCTATTACAAACAATGTTGAGAGCAGAGGATATACATPAAGACATCTTGA	912	
Qy	301	GIuYsLyAspGlySerIleYsProGluGluYsGluAspLySer	316	
Db	913	GAAAAGAAAGATGTTCAATAAAACCAAGAAAAAGAAATAAATCT	960	

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-Aug-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-900-963-38

Alignment Scores:
Pred. No.: 3,2e-122 Length: 1493
Score: 1465.00 Matches: 311
Percent Similarity: 98.42% Conservative: 0
Best Local Similarity: 98.42% Mismatches: 5
Query Match: 93.73% Indels: 3
DB: 3 Gaps: 0

US-09-837-344-31 (1-316) x US-09-900-963-38 (1-1493)
QY 1 SerAspLeuGluGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeu 20
DB 13 ACGGATCTAGAACCAAGAGAGACGCTCTAAAGAAAGTTGCCAAGAACCAAGACGATTTA 72
QY 21 GlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlu 40
DB 73 GAACCAAGATGACTTCTTAAGAAAGTTACCAAGACGACCAAGCGATTTAGAACCAAG 132
QY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnGlnArgAla 60
DB 133 AGACTTGTCTAA-GAAAGTTGCCAAGAACCAAGCGATCTTAAGAACCAAGAGACGCTGCT 191
QY 61 LysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArgArgAlaLysGluLys 80
DB 192 AAAAGAAAGTTGCCAAGAACCAAGCGATTTAGAACCAAGAGACGCTCTAAAGAAAG 251
QY 81 LeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlu 100
DB 252 TTGCAAGAACCAAGACGATTTAGAACCAAGATGACTTAAAGAAAGTTACCAAGAG 311
QY 101 GlnGlnSerAspLeuGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnSer 120
DB 312 CAGCAAGCGATTTAGAACCAAGAGACGCTTAAGAAAGTTGCCAAGAACCAAGAC 371
QY 121 AspLeuGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlu 140
DB 372 GATTTAGAACCAAGAGACGCTCTAA-GAAAGTTGCCAAGAACCAAGACGATTTAGAA 430
QY 141 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGln 160
DB 431 CAAGAGACGCTCTTAAGAAAGTTGCCAAGAACCAAGCGATTTAGAACCAAGAGAG 490
QY 161 ArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArgArgAlaLys 180

DB 491 CGTCTAAAGAAAGTTGCCAAGAACCAAGACGATTTAGAACCAAGAGACGCTCTAA- 549
QY 181 GlnLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArgArgAlaLysGluLysLeu 200
DB 550 GAAAGTTGCCAAGAACCAAGCGATTTAGAACCAAGAGACGCTCTTAAGAAAGTTG 609
QY 201 GlnGlnGlnGlnArgAspLeuGlnGlnGlnGlnLysAlaAspThrLysLysAsnLeuGluArg 220
DB 610 CAAGAGACGACCAAGAGATTTAGAACCAAGAGCGTGATACGAAAAAATTAGAAAGA 669
QY 221 LysLysGlnLysGlnAspLysLeuAlaGluAspLeuLysGlnArgLeuGlnLysProAla 240
DB 670 AAAAGAACCATGAGATATATTAGCAGAGGATTTATATGCTGTTGAAATACAGCT 729
QY 241 TleGluLeuProSerGluAsnGlnLysArgLysLysThrLysProHisGlnSerSerLeuPro 260
DB 730 ATAGAACTTCATCGAAGAAATGAAAGTGATATTTATATACCATCAATCTTTACCT 789
QY 261 GlnAspAsnArgGlnLysSerArgAspSerLysGlnLysSerLysLysGluLysThrAsn 280
DB 790 CAGGACCAACAGAGGAAATAGTAGATTCACAGGAAATATCTATATGAAAAAACAAT 849
QY 281 ArgGlnSerLysThrThrAsnValGluLysArgArgAlaLysLysGlnLysLeuGlu 300
DB 850 AGAGAACTCTATTACCAACAAATGTTGAAGACGAAAGGATATACATTAAGACATCTTGA 909
QY 301 GlnLysLysAspGlnSerLysLysProGlnGlnLysGlnAspLysSer 316
DB 910 GAAAGAAAGATGTTCAATTAACCAAGAACCAAGAAAGATTAATCT 957

RESULT 9
US-09-837-344-37
Sequence 37, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPIPODES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 954 base pairs

THIS PAGE BLANK (USPTO)

[illegible][illegible]

```
Db      826 AATGACTACTAGAAAAGAA----- 806
Qy      236 LeuGlulPrAlaIleGlulLeuProSerGluAenGluArgIYrTYrIleProHis 255
      : : : : :
Db      805 ---GATATCCCTGCTGTT-----TCTAAAAACGAGAA-----TTA 773
Qy      256 GluSerSerLeuProGlnAenAenArgIYAsnSerArgAenSerIYsgIuIleSerile 275
      : : : : :
Db      772 CAATCACCACAAACCA-----AACCTGACAGTAAA----- 743
Qy      276 IleGluYrThrAenArgIuSerileThrYrAen 287
      : : : : :
Db      742 ATAGAACCTACAACTTCAGAACTCTGTAATTTAAAT 707

RESULT 3
US-11-060-914-1
/ Sequence 1, Application US/11060914
/ Publication No. US20050261184A1
/ GENERAL INFORMATION:
/ APPLICANT: Kaufman, Paul L
/ APPLICANT: Geiger, Benjamin
/ APPLICANT: Bershadsky, Alexander
/ APPLICANT: Borras, Teresa
/ TITLE OF INVENTION: Method for Treating Glaucoma
/ FILE REFERENCE: 960296.00128
/ CURRENT APPLICATION NUMBER: US/11/060.914
/ CURRENT FILING DATE: 2005-02-18
/ PRIOR APPLICATION NUMBER: US 60/545,722
/ PRIOR FILING DATE: 2004-02-18
/ PRIOR APPLICATION NUMBER: US 60/545,723
/ PRIOR FILING DATE: 2004-02-18
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1
/ LENGTH: 5233
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (460)..(2838)
US-11-060-914-1

Alignment Scores:
Pred. No.: 1,49e-11 Length: 5233
Score: 241.00 Matches: 97
Percent Similarity: 48.73% Conservative: 76
Best Local Similarity: 27.32% Mismatch: 114
Query Match: 15.42% Indels: 69
DB: Gaps: 16

US-09-837-344-31 (1-316) x US-11-060-914-1 (1-5233)
Qy      3 LeuGlulGluIuAArgAlaIleYsgIuLeuGlnGlnGlnInserAenLeuGln 22
      : : : : :
Db      1273 TTGGAGCAGAGAAAGAA-----GAAAGATTAAAGCCGCAAGCAAAAGAAATAGCA 1326
Qy      23 AsparGleuAlaYsgIuYLeuGlnGlnGlnInserAenLeuGlnGlnIuArg 42
      : : : : :
Db      1327 GATAGCAGCAGAGAAATTGAGCAAGAAAGAGAGCTGCCCAAGAAAGAAAGAGGA 1386
Qy      43 AlaYsgIuYLeuGlnGlnGlnInserAenLeuGlnGlnIuArgAlaYsgIu 62
      : : : : :
Db      1387 GAGCAGAA-----GAGAGGGAAGATGAGGAGGAAGAAAGGAGCAGAG 1437
Qy      63 YLeuGlnGlnGlnInserAenLeuGlnGlnIuAArgAlaYsgIuYLeuGln 82
      : : : : :
Db      1438 -----GAGAGCAGAGATTAAGAGAGAAAGGAGCAGAG----- 1479
Qy      83 GluGlnInserAenLeuGlnGlnAenArgIuAlaYsgIuYLeuGlnGlnGln 102
      : : : : :
Db      1480 GAGAGGAGAGATTAAGAGAGAGAAAGGAGCAGAG-----GAGAGCAG 1530
Qy      103 SerAenLeuGlnGlnIuArgAlaYsgIuYLeuGlnGlnGlnInserAen 122
```

```
Db      1531 AGGATTAAGAGAGAAAGGAGCAGAG-----GAGAGGAGAAAGGCGCAGAGCA 1584
Qy      123 GluGlnGluIuArgAlaYsgIuYLeuGlnGlnGlnInserAenLeuGlnGln 142
      : : : : :
Db      1585 GAGAGAGAGAGAGCTTAAGTAAAGAGCAGAAAGCTTAACAGCAGCTTAAGAGAA 1644
Qy      143 ArgArgAlaYsgIuYLeuGlnGlnGlnInserAenLeuGlnGln 158
      : : : : :
Db      1645 AAACATGCTTACAGAGACAAAGATTAAGGAGGAGAAAGGTAGAACAGAAATGAAGG 1704
Qy      159 -----GluArgAlaYsgIuYLeuGln----- 167
Db      1705 AAATGGTAAATGAAGAAAGAGCAAGAGATTAACCTCAGACAGCTGCTTAAGAA 1764
Qy      168 -----GluGlnInserAenLeuGlnGlnIuArgAlaYsgIuYLeuGlnGln 185
      : : : : :
Db      1765 CAGGAGAGAGAGAGAAAGGAACTAAAGTCAGCAGCTTAAGAA-----GAAAGCTTCAGAA 1818
Qy      186 GluGlnInserAenLeuGlnGlnIuArgAlaYsgIuYLeuGlnGlnIuArg 205
      : : : : :
Db      1819 GACAGGCTTACCTCAAAAGAGAGATCAAGATGAAGATTAAGAAAGACAAAGAA 1878
Qy      206 AspleuGlnIuArgIYrYrIle-----AspThrIYrYrAen----- 217
      : : : : :
Db      1879 CCCAAAGAGAGAGTAAAGACCTTCATGATCGAAGAGGATTTACAGAAATTAACTG 1938
Qy      218 -----LeuGluArgIYrYsgIuIleGluYAsp-IlleuAlaGluAen 232
      : : : : :
Db      1939 CAGATGAGAAATTATGATGACCCCAAACTTAACATCTGAGATCTTCAGCGGCT 1998
Qy      232 YTYrGly-----ArgLeuGluIleProAlaIleGluLeuProse 245
      : : : : :
Db      1999 GAGAGGAGGCGCAGCGTGCACCCAGAGAGCTGAGGCGCCCGCCAGTGAAGCCGCG 2058
Qy      245 rGluAenGluArgIYrTYrIle-----ProHisInserSerLeuProG 261
      : : : : :
Db      2059 AAAAGCTGAGAGAGCTTCCTGCTGCTGCGGCGGAGACCGAGAGCAAGATTGAGAG 2118
Qy      261 naeAenArgIYAsnSerArgAenSerIYsgIuIleSerileIleGluYs-----Th 279
      : : : : :
Db      2119 CTCAAACAGAGCAGCAGAGGCG-GCTTGAGCTGAGAGAACTCAAGAAAGAGAGGA 2177
Qy      279 rAenArgIuSerileThrAenValGluYrArgArgAspIleHisYsgIuHisle 299
      : : : : :
Db      2178 GAGAGAGAGAGAGCTCTGAGAGAGAGAGCAGAGAGAG----- 2220
Qy      299 uGluGluYLeuAenArgIYserileYsProGluGlnIuYsgIu 313
      : : : : :
Db      2221 -CAGAGAGAGCCGATCGAAAACTCAGAGAGAGAGAGAGAG 2262

RESULT 4
US-10-955-054A-68
/ Sequence 68, Application US/10955054A
/ Publication No. US20050266420A1
/ GENERAL INFORMATION:
/ APPLICANT: PUSZTAI, LAJOS
/ APPLICANT: SYMMANS, W. FRASER
/ APPLICANT: HESS, KENNETH R.
/ APPLICANT: AYERS, MARK
/ APPLICANT: STEC, JAMES
/ TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY
/ FILE REFERENCE: UTXC.88005
/ CURRENT APPLICATION NUMBER: US/10/955.054A
/ CURRENT FILING DATE: 2004-09-30
/ NUMBER OF SEQ ID NOS: 195
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 68
/ LENGTH: 2617
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-955-054A-68
```


QY 134 GluGlnGlnSerAspLeuGluGlnGluArgAlaIysGluIleuGlnGln 153
|||
|||
Db 1960 TTGTTTAATTCAAGATGAATGATGGGTAAAGATGCTTTTGAAAGTTGAAAGACCA--- 2016
QY 154 SerAspLeuGluGlnGluArgAlaIysGluIleuGlnGlnGlnSerAspLeu 173
|||
|||
Db 2017 -----ATTGACGACCAAGAAATGTAAAGACAAATTAAATCAACATTAGTGAAGATT 2070
QY 174 GluGlnGlnArgArg---AlaIysGluIleuGlnGlnGlnSerAspLeuGlu--- 191
|||
|||
Db 2071 AAACAGCAACTTAAGATCTTAATGAGAAATCGAAATTAATGAAGTCAAGCTTCAAAA 2130
QY 192 -----GlnGluArgArgAlaIysGlu--- 198
|||
|||
Db 2131 TGTGATCAAGATATACCTTTCTATAGAAATCATTTATCAAGATTAATGAACAAATCA 2190
QY 199 -----LysLeuGln 201
|||
|||
Db 2191 AAGCTAGATGATTAAATCAACCATGCAATGATCATTTAAATGACACGTATCAACTCA 2250
QY 202 GluGlnGlnArgAspLeuGluGlnArgIleAlaAspThrIleIysAsnLeuGluArgIys 221
|||
|||
Db 2251 GTAGAACTGCAAGATGCAATGATCTGATGAAACTATTGACAAATTTCCGTAAGAAA 2310
QY 222 LysGluIleGluAspIleLeuAlaGluAspLeuTyrgIleArgLeuGluIleProAlaIle 241
|||
|||
Db 2311 GTAAATTTAAACGAATGACATGCAATGATTA---GGTCTGTAATTTAAATGACAAATA 2367
QY 242 GluLeuProSerGlu---AsnGluArgGlyTyrgIleProHisGlnSerSerLeuPro 260
|||
|||
Db 2368 GAACAATTGAAAGCTTGAATGAACGATACACATTTTAAATGAGCAA----- 2415
QY 261 GlnAspAsnArgGlyAsnSerArgAspSerIleGluIle-----SerIleGluIys 278
|||
|||
Db 2416 -----CGAACGATTTAAGAGAACAAAGAACCTTAGAACAAATCATTCATGAA 2466
QY 279 ThrAsnArgGluSerIleThrThrAsnValGluGlyArg 291
|||
|||
Db 2467 ATGATTAAGAA-----GTTGAAGTCTGT 2490

RESULT 6
US-10-793-626-4160
; Sequence 4160, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4160
; LENGTH: 3520
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4160

Alignment Scores:
Pred. No.: 3,286-11 Length: 3520
Score: 235.00 Matches: 89
Percent Similarity: 49.58% Conservative: 86
Best Local Similarity: 25.21% Mismatches: 90
Query Match: 15.04% Indels: 88
DB: 6 Gaps: 17

US-09-837-344-31 (1-316) x US-10-793-626-4160 (1-3520)

QY 4 GluGlnGluArgArgAlaIysGluIleuGlnGln----- 16
|||
|||
Db 2366 GAACAAAAAACAAAGCT---GAACAAATTAAAGTAAATTTTAACTGCAAGTCAAGCAG 2422
QY 17 -----GlnSerAspLeuGluGlnGlnAspArgLeuAlaIysGluIleuGlnGln 32
|||
|||
Db 2423 TACAACAATTTAAAGAAACAGTACATCATCAGAAATTTAAAGCTGATACACTTAAACA 2482
QY 33 GlnGlnSerAspLeuGluGlnGlnGlnArgAlaIysGluIleuGlnGlnGlnSer 52
|||
|||
Db 2483 CAAGAACACATCTTTAAATGAACAT-----GAAGAGTTGAAATTTGAAAAAAT 2533
QY 53 AspLeuGlnGlnGlnGlnGlnArgAlaIysGluIleuGlnGlnGlnGlnGlnGlnGln 72
|||
|||
Db 2534 GATGGAATATCAAGATGATAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2593
QY 73 GlnGluArgArgAlaIysGluIleuGlnGlnGlnSerAspLeuGlnGln---Asp 91
|||
|||
Db 2594 GAG-----ATACAAACAATTTGACCACTAGAAAGTATTTGAAAGTATACA 2644
QY 92 ArgLeuAlaIysGlu-----LysLeuGlnGln 101
|||
|||
Db 2645 CAATTATCAAAAGAAAGAAAGCTTGCACATCAACACAAACAACTTACATCAAAAA 2704
QY 102 GlnSerAspLeu-----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 114
|||
|||
Db 2705 CAATCTGATTTAGCTGTGTTAAAGACGATTAACGAAAGCAAGTTTATGAAAGCT 2764
QY 115 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 133
|||
|||
Db 2765 TTGATTAACAACTTACGATTCGAAACGTCAAATAATGAAAGTAAAGTAAAGTAAAGTAA 2824
QY 134 GlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 153
|||
|||
Db 2825 TTGTTTAATTCAGATGAATGATGCGTAAAGATGCTTTGAAAGTGAAGAACAA--- 2881
QY 154 SerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 173
|||
|||
Db 2882 -----ATTGACGACCAAGAAATGTAAAGCAAAATTTAAATCAACACTTAGGAAGATT 2935
QY 174 GluGlnGlnArgArg---AlaIysGluIleuGlnGlnGlnGlnGlnGlnGlnGlnGln 191
|||
|||
Db 2936 AAACAGCAAGCTAAAGATCTTAATGAGAAATGCAATTAATGAAGTCAAGCTTCAAAA 2995
QY 192 -----GlnGluArgArgAlaIysGlu--- 198
|||
|||
Db 2996 TGTGATCAAGATATACCTTTCTATAGAAATCATTCAGATATTAAGCAAAAAAATCA 3055
QY 199 -----LysLeuGln 201
|||
|||
Db 3056 AAGCTAGATGATTAAATCAACCATGCAATGATCATTTAAATGACACGTATCAACTACA 3115
QY 202 GluGlnGlnArgAspLeuGlnGlnArgIleAlaAspThrIleIysAsnLeuGluArgIys 221
|||
|||
Db 3116 GTAGAACTGCAAGATGCAATGATCTGATGAAACTATTGCAATTTTCCGTAAGAAA 3175
QY 222 LysGluIleGluAspIleLeuAlaGluAspLeuTyrgIleArgLeuGluIleProAlaIle 241
|||
|||
Db 3176 GTAAATTTAAACGAATGACATGATGAATTA---GGTCTGTAATTTAAATGACATA 3232
QY 242 GluLeuProSerGlu---AsnGluArgGlyTyrgIleProHisGlnSerSerLeuPro 260
|||
|||
Db 3233 GAACAATTTGAAAGTGAATGAACGATACACATTTTAAATGACAAA----- 3280
QY 261 GlnAspAsnArgGlyAsnSerArgAspSerIleGluIle-----SerIleGluIys 278
|||
|||
Db 3281 -----CGAACGATTTAAGAGAACAAAGAACCTTAGAACAAATCATTCATGAA 3331
QY 279 ThrAsnArgGluSerIleThrThrAsnValGluGlyArg 291
|||
|||
Db 3332 ATGATTAAGAA-----GTTGAAGTCTGT 3355

Db 421 AAGAGAAAAGAGAACTAGGATCTGATCCATTCAGCAGAAAGGGGGAGCATCTCAGA 480
Qy 45 Gtlybyleuglnglnglnglnserapleuglnglnglnuargalalybgluyleu 64
Db 481 GTGAAAGAAAAGAAAGA-----GAGCAGCAAGCTCGGAAACGACATCGAGAA 528
Qy 65 Glnlglnglnglnserapleuglnglnglnuargalalybgluyleuglnglngln 84
Db 529 GAACAGATTAAGCTCCCGGGAATGGGAAGACAGAAAGAGAAATGCGAAGGAG 588
Qy 85 Glnserapleuglnglnglnserapleuialalybgluyleuglnglnglnglnserap 104
Db 589 CATTCAGAGAGAAAGGACCGCTTG-----GAGCATTTAGAAAGAACCG----- 636
Qy 105 Leuglnglnglnuargalalybgluyleuglnglnglnglnserapleuglngln 124
Db 637 ---GAGCGGAGCGGAGATGCGGAGCAGCAAGAGAGCAGCGGACAGAGAGAGCGC 693
Qy 125 Glnuargalalybgluyleuglnglnglnglnserapleugln----- 140
Db 694 GAGCGCGGCGGAGAGCGCGCAGAGAGCGGAGCGCGAGGAGTGTCTGCACAT 753
Qy 140 ----- 140
Db 754 CACCGAAGATGAGAGAGACTACAGCAGCAAAAGTGAAGCCAGCCACTGAGATCGCAGC 813
Qy 141 -----Glnglnuarg----- 143
Db 814 CCGCTCGCGCGCGCGGAGCGGTTCAGTTGGAGACGGCGGAGCCAGGTGAGGCC 873
Qy 144 -----ArgAla-----LysGluyleuglnglnglngln 153
Db 874 AGCGCGCGCGCGCGCAGAGCAGCAGCTTAAAGAAAGAAATGAAAGAAAGGAGC 933
Qy 154 -----Serapleuglnglnglnglnuargalalybgluyleuglnglnglnglnser 171
Db 934 CTGCTGCTCGGACTTTCAGAGACTTCAGCAGCAGCGAGAGAAAGCAAGCTCGGCGGAGTCC 993
Qy 172 Apleuglnglnglnuargalalybgluyleuglnglnglnglnserapleugln 191
Db 994 TCGTCAGAGATCAGCGCTTCAGGAGAAAGAGAGAGAGAGAGAGAGAGAG 1053
Qy 192 Glnlgluargalalybgluyleuglnglnglnglnuargapleuglnglnuarglys 211
Db 1054 GAGGAAGGAGAGCAAGTGAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1113
Qy 212 AlaapThrlyblybapleuglnglnuarglyblybgluyleu-----AspIleuAla 229
Db 1114 GAGGAGAGCGGAGCACTTCAGAGAGCATCAGACGCTCGCCGAAGAAAGTAAAGTAG 1173
Qy 230 GlnapleuglyblybargleuglnglnleproAlaIleGluLeuProSerGluasnGluarg 249
Db 1174 GAGGAAGATG-----AGTGAAGATGAAGAAAGCAAGAAATGATAAAC 1212
Qy 250 GtlytyrTyleproHleGlnserleuProGlnaapapnarglyblybserargasp 269
Db 1213 CACCTCTTGAGTTCTTCAGAGATCAGCGTTTCAGCCAGATTC-----GGGAGAGTGAAGAA 1269
Qy 270 SerlybgluIleSerIleIleGluyleThraaargGluSerIleThrThr 286
Db 1270 GCAGAGAA---GAAGTGGGTGAGGAAGCGCGCAGAGAGCGCCCTGACA 1317

RESULT 9
US-10-793-626-3277/c
; Sequence 3277, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793, 626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3277
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3277

Alignment Scores:
Pred. No.: 2,86e-11 Length: 519
Score: 225.50 Matches: 64
Percent Similarity: 59.89% Conservative: 42
Best Local Similarity: 36.16% Mismatches: 62
Query Match: 14.43% Indels: 9
DB: 6 Gaps: 8

US-09-837-344-31 (1-316) x US-10-793-626-3277 (1-519)

Qy 27 LysGluyleuglnglnglnglnserapleugln---Glnlgluargalalybglu 45
Db 510 AAGCAGAAAGAGAGAGCAATGCAAGACCACTCAATCTGAGCCAGCAAGAGAGAGAG 451
Qy 46 Lysleuglnglnglnglnserapleuglnglnglnuargalalybgluyleugln 65
Db 450 GAGCAATGCAAGAGAGCAATCTGAGCCAGCAAGAGAGAGAGAGAGAGAGAGAG 391
Qy 66 Glnlnglnserapleuglnglnglnuargalalybgluyleuglnglnglngln 85
Db 390 ---AAGACCTCAATCTGAGCCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 337
Qy 86 Serapleuglnglnglnuargalalybgluyleuglnglnglnglnserapleu 105
Db 336 CTCAATCTGAGCCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 283
Qy 106 Glnlnglnuargalalybgluyleuglnglnglnglnserapleuglngln 125
Db 282 CTGAGCCAGCAAG 223
Qy 126 ArgargalalybgluyleuglnglnglnglnserapleuglnglnglnuargAla 145
Db 222 AGGCAAG 166
Qy 146 LysGluyleuglnglnglnglnserapleuglnglnglnuargalalybgluyle 165
Db 165 AAGGAG 109
Qy 166 Leuglnglnglnglnserapleuglnglnglnuargalalybgluyleuglngln 185
Db 108 ATGACAG---AAGCAGCTCAATCTGAGCCAGCAAGAGAGAGAGAGAGAGAGAGAG 55
Qy 186 Glnlnglnserapleuglnglnglnuargalalybgluyleuglngln 202
Db 54 AAGCACTCAATCTGAGCCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4

RESULT 10
US-10-623-155-119
; Sequence 119, Application US/10623155
; Publication No. US20050261166A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongfong
; APPLICANT: Peckham, David W.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C20
; CURRENT APPLICATION NUMBER: US/10/623, 155
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 560

```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119
; LENGTH: 8948
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-155-119

Alignment Scores:
Pred. No.: 1,03e-08 Length: 8948
Score: 211.50 Matches: 88
Percent Similarity: 42.18% Conservative: 71
Best Local Similarity: 23.34% Mismatches: 131
Query Match: 13.53% Indels: 87
DB: Gaps: 13

US-09-837-344-31 (1-316) x US-10-623-155-119 (1-8948)

QY 2 AapLeuGluGlnGluArg-----ArgAlaLysGluLysLeuGln 14
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5226 GACCTCCGCGCAGAGGACGCTGCGATGCCACCTGAGGAAAGCAGACCCAG 5285

QY 15 GlnGlnGlnSerAapLeu-----GlnGlnAapArgLeuAlaLysGluLysLeuGln 32
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5286 GAAGAGCTGAGAGGAGCTCTCTCTGAGTCGAGCCCTGAGCGCGAGTTACTCCAGGA 5345

QY 33 Gln-----GlnSerAapLeuGlnGlnLysArgAlaLys----- 44
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5346 CAGGAAAGTGTCAAAAGCTCCTGAGGATGACATTTCCAGAAAGCGGATAGAAAGT 5405

QY 45 -----GluLysLeuGlnGlnGlnGlnSerAapLeuGlnGlnLysArgAlaLysGlu 62
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5406 AAAAGCAGAACTTAATGAACCAAAATGAATTAAGAGCGTGAGCTCTCACAGAG 5465

QY 63 LysLeuGlnGlnGlnGlnGlnSerAapLeuGlnGln-----ArgAlaLys 78
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5466 AACCTGACCAAGAGAGCATTTGATGTAAGAAAGAACTGCGGAACCTGAGCTGAGTAC 5525

QY 79 GluLysLeuGlnGlnGlnGlnGlnSerAapLeuGlnGlnAapArgLeuAlaLysGluLysLeu 98
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5526 GATGACCTGAGAGAGAGCAGACCAAGCGGACAGTATGAATGA-----ACCATC 5579

QY 99 GlnGlnGlnGlnSerAapLeuGlnGlnGlnLysArgAlaLysGluLysLeuGlnGln 118
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5580 TTGGAACTAAGAGACCGCTGCAATCAGCAACACCGGACCTGGAATCGCAGGGGCTG 5639

QY 119 GlnSerAapLeuGlnGlnGlnGln----- 126
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5640 ATTAATGATTTACAGAGAGAGGAAATTTGAGACAGGAAATTGCAAAATTCAGAAAG 5699

QY 127 -----ArgAlaLysGluLysLeuGlnGlnGlnGlnSerAapLeuGlnGln 143
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5700 CAGGCTTTAGAGGATCTTAATAGGATTCAGGAATCAAGATCACTGACTGAGTGCTA 5759

QY 144 ArgAlaLysGluLysLeuGlnGlnGlnGlnGlnSerAapLeuGlnGlnGlnGln 160
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5760 CAGAAAGAGAGAGACCTTCTGCGTGAATAATCAAGCTCTGAGCAAGCAAGGAGCTG 5819

QY 161 ---ArgAlaLysGluLysLeuGlnGlnGlnGlnGlnSerAapLeuGlnGlnGlnGln 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5820 CAGAGGCTGAGAGTGAAGTGAATTCGCAAAATCACTTACAGGAGCAAGCAAGGAGCTG 5879

QY 180 LysGluLysLeu-----GlnGlnGlnGlnSerAapLeuGlnGln----- 192
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5880 AAACAGGCGCTGAGAGTGAAGAAACAGCAAAATTCAGAAATCACTGAGTGAAGACT 5939

QY 193 -----GluArgArgAlaLysGluLysLeuGlnGlnGlnGlnGlnAapLeuGln 208
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5940 CAATATTCCCGCAAGAGAGAGGCTATTAGAAAGTAAATTCGAAAGAGAAAGAGAGAG 5999

QY 209 GlnArgLysAlaAapThrLysLysAapLeuGlnGlnGlnGlnGlnGlnGlnGln 228
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6000 AGAGAGAAAGAAAGCTTTAGAGAGTGAAGTGAAGAAAGCTCAAGAGAGATCAAGAGATT 6059
```

```
QY 229 AlaGluAapLeuArgLysLeuGlnLeuProAlaIleGluLeuProSer-----Glu 246
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6060 GAAGAGAGTGCAGGCGGTAAGCTGAGATTCTACAGAGAGACACAGTACAGTTAGAA 6119

QY 247 AapGluArgGlyTyrTyrTyrIleProHisGlnSerSerLeuProGlnAapAapArgLysAap 266
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6120 ACAGAAACCTCCCGATATCAGAGGAGATGTAATTAATCAAGACAGCCCATATAGGCTC 6179

QY 267 SerArgAapSerLys----- 271
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6180 CATCGAGAGACCCACAGACTGAGTGTGATGAGACCGTTGACACCTCCAGCTGCTTTGAT 6239

QY 272 -----GluIleSerIleIleGluLysThr 279
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6240 GGGCTGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6299

QY 280 AapAapGluSerIleThrThrAapValGluLysArgAapIleHisLys 296
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6300 ACCTTGACAAACTA-----TTGAAGGGAAGAAAGTCAAGTGAAGAA 6341

RESULT 11
US-10-821-234-198/c
; Sequence 198, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Labat, Ivan
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 198
; LENGTH: 6906
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-198

Alignment Scores:
Pred. No.: 1,17e-08 Length: 6906
Score: 209.50 Matches: 83
Percent Similarity: 46.60% Conservative: 54
Best Local Similarity: 28.23% Mismatches: 106
Query Match: 13.40% Indels: 51
DB: Gaps: 13

US-09-837-344-31 (1-316) x US-10-821-234-198 (1-6906)

QY 3 LeuGlnGlnLysArgArgAlaLysGluLysLeuGlnGln----- 17
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3633 CTGAGAGAGAGCGGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3574

QY 18 -----SerAapLeuGlnGlnAapArgLeuAlaLysGlu 28
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3573 ACTGACTCTTAACCTCTGCTTCTCCCTGACTGCAAGAGAGTGAAGCCCGAA 3514

QY 29 Lys-----LeuGlnGlnGlnGlnSerAapLeuGlnGlnGlnGlnGlnGln 45
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3513 AACGCTTCACACTCCAGCGGCGCACAGAAACAGTATTCGGGAG----- 3469

QY 46 LysLeuGlnGlnGlnGlnGlnSerAapLeuGlnGlnGlnGlnGlnGlnGlnGln 65
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3468 ---CTGCAAGCTTCAGAGAGAGCGCCGCAAGTCAAGCGGAGAGAGAGAGAGAGAGAG 3412

QY 66 GluGlnGlnSerAapLeuGlnGln-----GluArgArg 76
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3411 GTCCAGCAAGAGAGCTTCTTCGGGAGAGTCTGCCCCCGGACCCGTGAAGCGGAGAC 3352
```



```
QY      241 IIEGIuleuProSerGIuAnGIuArgGIyTYrTYrIleProHISGInSerSerLeuPro 260
      ::: ||||| ||||| ||||| |||||
Db      3986 GTGGCTGCAGCATCTCAAGAGAAAG-----CCCAAGCTGCTGAGACACCC 4033
QY      261 GInAspAsnArgGIy-AsnSerArgAsp--SerLySGluIleSerIleIleGIuLySTh 279
      ::: ||||| ||||| ||||| |||||
Db      4034 GCCAGCTCCCGGAGAGCCCTCTCGGACCTGCCTCAAGTTGAGGGAGAGCCGAGGAGAC 4093
QY      279 rAsnArg-----GluSerIleThrThrAsnValGIuGI 290
      ||||| ||||| ||||| |||||
Db      4094 GCAGAGCACACTGCAGGCCGAGTGTACCACTACCGCAGCATCTGCGGAGACGAGGG 4153
QY      290 yArg--ArgAspIleHisLySGIyHisLeuGIuGIuLyS-----LyAsAs 304
      ||||| ||||| ||||| |||||
Db      4154 CATGCTCAAGACCTGCAGAAAGCCTGGAGAGAGAGAGACAGCTGTGAGGGCCAAAGT 4213
QY      304 pGIySerIleLySProGIuGIuLySGIuAspLyS 315
      ||||| ||||| ||||| |||||
Db      4214 GGGCGCGGAGAGAGAGCTCCAGAAATCCCGG 4247
```

Search completed: December 7, 2005, 16:25:08
Job time : 443.478 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:12:18 ; Search time 24.0831 Seconds
(without alignments)
1262.481 Million cell updates/sec

Title: US-09-837-344-31

Perfect score: 1563
Sequence: 1 SDLEQERRAKKLEQEQSD...GHLEKKDSIKPEQKEDKS 316

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1494	95.6	1909	2	A45592
2	328	21.0	1407	1	S28589
3	318.5	20.4	737	2	T15597
4	318.5	20.4	1017	2	T15598
5	317.5	20.3	1898	1	A45973
6	310.5	19.9	1549	1	A40691
7	289	18.5	407	2	S23335
8	287.5	18.4	1432	2	B85431
9	283.5	18.1	451	2	G70241
10	272	17.4	699	2	B84555
11	271.5	17.4	887	2	G88464
12	269	17.2	402	2	S37046
13	269	17.2	1979	1	S03166
14	267.5	17.1	729	2	T50599
15	266	17.0	877	2	T50591
16	261.5	16.7	649	2	T72732
17	261	16.7	746	2	T72737
18	260.5	16.7	1166	2	H86341
19	260	16.6	771	1	A23430
20	259	16.6	1840	2	T29091
21	258.5	16.5	792	2	T42963
22	258	16.5	386	2	S05568
23	255.5	16.3	501	2	C71948
24	254.5	16.3	1961	1	A61231
25	254.5	16.3	1963	1	MMKW
26	253.5	16.2	520	2	S35575
27	253.5	16.2	1039	2	S18199
28	253	16.2	1738	2	T14867
29	253	16.2	810	2	T44430

30	253	16.2	1940	2	A59287	myosin heavy chain
31	253	16.2	2101	2	A42184	nuclear mitotic ap
32	252.5	16.2	390	2	T34137	hypothetical prote
33	252.5	16.2	1959	1	A33877	myosin heavy chain
34	252.5	16.2	1964	2	A59282	nonmuscle myosin I
35	252	16.1	1701	2	T09127	probable erythrocy
36	251.5	16.1	377	2	S52537	emm 115 protein -
37	251.5	16.1	384	2	S49550	M-like protein emm
38	251.5	16.1	1110	2	IS1116	NP-180 - sea lampy
39	251	16.1	839	2	IS0590	class I INCENP pro
40	250.5	16.0	1976	2	A59252	myosin heavy chain
41	250	16.0	2007	1	B43402	myosin heavy chain
42	249.5	16.0	347	2	I46592	involucrin - pig
43	249.5	16.0	465	2	A02986	myosin alpha heavy
44	249.5	16.0	527	2	S33068	myosin heavy chain
45	249.5	16.0	1085	2	S62516	hypothetical coile

ALIGNMENTS

RESULT 1

A45592
liver stage antigen LSA-1 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C>Date: 22-Nov-1993 #sequence revision 02-Dec-1994 #text change 09-Jul-2004

C/Accession: S24597; A45592; S29393; S34842; B45592; C45592; D45592

R/Zhu, J.; Hollingdale, M.

Submitted to the EMBL Data Library, November 1990

A/Reference number: S24597

A/Accession: S24597

A/Molecule type: DNA

A/Residues: 1-1909 <ZHU>

A/Cross-references: UNIPROT:Q25893; UNIPARC:UPI0000081753; EMBL:X56203; NID:99915; PID:

R/Zhu, J.; Hollingdale, M.R.

Mol. Biochem. Parasitol. 48, 223-226, 1991

A/Title: Structure of Plasmodium falciparum liver stage antigen-1.

A/Reference number: A45592; PMID:92107224; PMID:1640628

A/Accession: A45592

A/Molecule type: DNA

A/Residues: 1-195,638-688;1165-1215;1590-1909 <ZH2>

A/Cross-references: UNIPARC:UPI0000177378; UNIPARC:UPI0000177379; UNIPARC:UPI000017737A

A/Title: A liver-stage-specific antigen of Plasmodium falciparum characterized by gene

A/Reference number: S29393; PMID:87315391; PMID:3306406

A/Accession: S29393

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 333-387 <GUE1>

A/Cross-references: UNIPARC:UPI000017737C; EMBL:M28266

R/Guerin-Marchand, C.; Drulhe, P.; Galey, B.; Londono, A.; Patartopitkul, J.; Beaudoin

Submitted to the EMBL Data Library, April 1992

A/Description: a liver-stage-specific antigen of Plasmodium falciparum characterized by

A/Reference number: S34842

A/Accession: S34842

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 333-381, 'HKA1' <GUE2>

A/Cross-references: UNIPARC:UPI000017737D; EMBL:M28266

A/Note: difference at carboxyl end due to frameshift error

C/Comment: This protein is found as flocculent material in the parasitophorous vacuole.

C/Superfamily: trichohyalin; calmodulin repeat homology

C/Keywords: EF hand

F:154-1629/Region: 17-residue repeats (A-K-E-K-L-Q-E-Q-S-D-L-E-Q-E-R-R)

Query Match 95.6%; Score 1494; DB 2; Length 1909;
Best Local Similarity 95.9%; Pred. No. 3.3e-61;
Matches 303; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
1 SDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRA 60
|||||


```

Oy 63 KLO-EOGSDLEOERRAKKLOEQOSDLEODRLAKKLOEQOSDLEOERRAKKLOEQOSD 121
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 293 RIETIERIERIKREIERERERERERERERERERERERERERERERERERERERERER 350
Oy 122 L-----EOERRAKKLOEQOSDLE-----EOERR-----AKKLOEQOSDLEOERRA 162
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 351 LORREADEDRERORLEDEARRERERERERERERERERERERERERERERERERERERER 410
Oy 163 KKKLOEQOSDLEQ-----ERRAKKLOEQOSDLE-----QERRAKKLOEQOSDLEOERRK 211
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 411 RIRLEOQKIDMERIDARRERERERERERERERERERERERERERERERERERERERER 469
Oy 212 ADTKKLERKKEHGDILAEIDLPALEPSENE-----RCGYIIPHOSSLPQDNKGN 266
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 470 EEEERERERERERERERERERERERERERERERERERERERERERERERERERERER 529
Oy 267 SROSKESITIEKTNRESITTNVEGRDINKHGLE 300
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 530 ELDRKAQELSEREMRE-----KERRRERANEE 557

RESULT 5
A45973
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C:Accession: A45973
R:Lee, S.C.; Kim, I.G.; Markov, L.N.; O'Keefe, E.J.; Parry, D.A.D.; Steinert, P.M.
J. Biol. Chem. 266, 12164-12176, 1993
A:Title: The structure of human trichohyalin. Potential multiple roles as a functional E
ed (cross-linking) protein.
A:Reference number: A45973; MUID:93280194; PMID:7685034
A:Molecule type: DNA
A:Residues: 1-1898 <LE>
A:Cross-references: UNIPROT:Q07283; UNIPARC:UPI000013738B; GB:U09190; NID:9292835; PIDN:
A:Note: authors translated the codon AGG for residue 1714 as Pro
C:Comment: Trichohyalin is a protein of the medulla of the hair and of the inner root sh
C:Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat
C:Genetics:
A:Gene: GDB:THH
A:Cross-references: GDB:136223; OMIM:190370
A:Map position: 1q21-1q21
C:Superfamily: trichohyalin; calmodulin repeat homology
C:Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat
F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 20.3%; Score 317.5; DB 1; Length 1898;
Best Local Similarity 29.7%; Pred. No. 1.1e-07;
Matches 105; Conservative 77; Mismatches 116; Indels 55; Gaps 14;

Oy 4 EOERRAKKLOEQOSDLEODRLAKKLOEQOSDLEOERRAKKLOEQOSDLE-----E 55
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 458 EOERR-EOURLKREOE-RRDWLKRBEETERN---EOERRKQQLKRDQEEERERERERER 512
Oy 56 QERRAKKLOEQOSDLEOERRAKKLOEQOSDLEODRLAKKLOEQOSDLEOERRAKKLOEQ 115
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 513 EERREOQERRRQQLRBEERERERERERERERERERERERERERERERERERERERERER 569
Oy 116 -OEOQSDLEOERRAKKLOEQOSDLEOERRAKKLOEQOSDLEOERRAKKLOEQOSD-L 173
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 570 KRBEERKLEOERRERERERERERERERERERERERERERERERERERERERERERER 629
Oy 174 EOERRAKKLOEQO-----SDLEOERRAKKLOEQOORDLEOR--KADTKKLER 220
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 630 EOERRRDRRLKREBEERERERERERERERERERERERERERERERERERERERERER 689
Oy 221 --KKHGDILAEIDLPALEPSENE-----IPALEPSENERGYIIPHOSSL-----PODN 263
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 690 RLKREHBEERERERERERERERERERERERERERERERERERERERERERERERERER 746
Oy 264 RGNBRDSKESITIEKTNRESITTNVEGRDINKHGLEKKGOSIKPEQKEDKS 316
```

```

Db 747 RAAPQEOE-----EKRRRESLEOQEEERERERERERERERERERERERERERERER 795

RESULT 6
A40691
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 21-Sep-1993 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C:Accession: A40691; A34209; S32633
R:Fieltz, M.J.; McLaughlan, C.J.; Campbell, M.T.; Rogers, G.E.
J. Cell Biol. 121, 855-865, 1993
A:Title: Analysis of the sheep trichohyalin gene: potential structural and calcium-bind
A:Reference number: A40691; MUID:93260018; PMID:7684041
A:Accession: A40691
A:Molecule type: DNA
A:Residues: 1-1549 <FIE>
A:Cross-references: UNIPROT:P22793; UNIPARC:UPI000013738D; EMBL:Z18361; NID:9295940; PI
R:Fieltz, M.J.; Presland, R.B.; Rogers, G.E.
J. Cell Biol. 110, 427-436, 1990
A:Title: The cDNA-decoded amino acid sequence for trichohyalin, a differentiation marke
A:Reference number: A34209; MUID:90130632; PMID:2298812
A:Accession: A34209
A:Molecule type: mRNA
A:Residues: 1016-1151, 1205-1257, 1281-1398, 'G', 1400-1549 <F12>
A:Cross-references: UNIPARC:UPI000016C5F4; GB:X51695; NID:91827; PIDN:CAJ5992.1; PID:G
C:Comment: Trichohyalin is a protein of the medulla of the hair and of the inner root sh
C:Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat
F:49-81/Domain: calmodulin repeat homology <EF2>
F:387-891/Region: 28-residue repeats
F:886-1519/Region: 23-residue repeats

Query Match 19.9%; Score 310.5; DB 1; Length 1549;
Best Local Similarity 27.7%; Pred. No. 2e-07;
Matches 105; Conservative 79; Mismatches 122; Indels 73; Gaps 14;

Oy 3 LEORRAKKEKLOEQOSDLE-----EOURLAKKLOEQOSDLEOERRAKKLOEQOSDLEOE 57
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 203 VEOQQRGRGRLRERERERERERERERERERERERERERERERERERERERERERERER 262
Oy 58 RRAKKEKLOEQOSDLEOERRAKKLOEQO-----OSDLEODRLAKK-----LOEQO 102
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 263 RKKERERERERERERERERERERERERERERERERERERERERERERERERERERER 322
Oy 103 S-----DLEORRAKKEKLOEQOSD-----LEORRAKKEKLOEQOSDLEOERR-----A 145
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 323 SRROEQELLRROEQOQISEEVQSLQOEDQGRORLQOEDQRYQDNRMWQLEESQORRYTLVA 382
Oy 146 KKKLOEQOSDLEOERRAKKLOEQOSDLEOERRAKKLOEQOSDLEOERRAKKLOEQOSDLE 197
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 383 KPAQREYVREBEQURLKEKLEQRRERERERERERERERERERERERERERERERERERER 442
Oy 198 EKLQEOORD-LE-----ORKADTKKLERKK--EKGDILAEIDLPALEIPALE 243
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 443 RRQERERERERERERERERERERERERERERERERERERERERERERERERERERER 500
Oy 244 PSENERGY-----IPHOSSLPQDNKGNRDSKESITIEKT-----NRESITTNVEGRDINK 296
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 501 ROERERQYLKVELQOBEQOQREERERERERERERERERERERERERERERERERERERER 555
Oy 297 GHLEKQDSIKPEQKEDK 315
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 556 RROERERQYLEKVELQOBE 574
```

RESULT 7
S23325

M2 protein precursor - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
A:Variatey: serotype M2
C:Date: 22-Nov-1993 #sequence revision 10-Nov-1995 #text_change 09-Jul-2004
A:Accession: S23325; S35761; S61078; S60785
R:Bessen, D.E.; Fischetti, V.A.
Infect. Immun. 60, 124-135, 1992
A:Title: Nucleotide sequences of two adjacent M or M-like protein genes of group A streptococci
A:Reference number: S23325; MUID:92104662; PMID:1370269
A:Accession: S23325
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-407 <BES>
A:Cross-references: UNIPROT:P50468; UNIPARC:UPI000012EABE; EMBL:X61276; NID:g47369; PIDD:R.Podbielski, A.
submitted to the EMBL Data Library, November 1992
A:Reference number: S35760
A:Accession: S35761
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-17 <POD>
A:Cross-references: UNIPARC:UPI000017823B; EMBL:X69324
R:Whitmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kenoe, M.
Submitted to the EMBL Data Library, July 1994
A:Description: Noncongruent relationships between variation in emm1 gene sequences and t
A:Reference number: S61072
A:Accession: S61078
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 12-94 <MHA>
A:Cross-references: UNIPARC:UPI0000081BF6; EMBL:U11958; NID:g533593; PIDD:AAA99574.1; P
R:Whitmore, A.M.; Kapur, V.; Sullivan, D.D.; Musser, J.M.; Kenoe, M.A.
Mol. Microbiol. 14, 619-631, 1994
A:Title: Non-congruent relationships between variation in emm gene sequences and the pop
A:Reference number: S60784; MUID:95198537; PMID:7891551
A:Accession: S60785
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 31-89 <MH2>
A:Cross-references: UNIPARC:UPI000017823C; EMBL:U11958
C:Superfamily: M5 protein

Query Match 18.5%; Score 289; DB 2; Length 407;
Best Local Similarity 29.6%; Pred. No. 6.2e-07;
Matches 89; Conservative 69; Mismatches 105; Indels 38; Gaps 12;

QY 4 EERRAKETLQEQOSDLEQ-----DLRAKETLQEQOSDLEGERAKETLQEQOSDLEQ 57
DB 90 EHHKHDKHELEKSEVENHYRLQDQYKEQ-QEQQNLEF-----LERQGRVE 140
QY 58 RRAKETLQEQOSDLEGERAKETLQEQ-OSDLEQRLAKETLQEQOSDLEGERAKETLQ 116
DB 141 KKYQELQYQ-QLEKEKISSEASRKSRLRDLSEASRAKKDLFAHQKLEKEKQISEASR 199
QY 117 EQ-QSDLEGERAKETLQEQOSDLEGERAKETLQEQOS-DLEGERAKETLQEQOSDLE 174
DB 200 KSLRDLSEASRAKKDLFAHQKLEKEKQISEASRQGSRLDLSEASRAKKDLFAHQKLE 259
QY 175 QERRAKETLQEQOS-DLEGERAKETLQEQOSDLEQ-RKADTKKNLEKKKHGDI LAE- 230
DB 260 EEKQISEASRQGSRLDLSEAREKKV--EADLEANSKLTALETLNKELEBGGKLSBK 316
QY 231 ---DLGRLEIPALIEPSE-----NERGYIIPHQSLLPODNRGNSRDSKEISILEK 278
DB 317 EKRAELQAKLEAKKALKEQLAKQAEELATLKNQPTNANVAQANRSRQAMTQQRITLPS 376
QY 279 T 279
DB 377 T 377

RESULT 8
B85431

[illegible]

Best Local Similarity 28.2%; Pred. No. 1.2e-06;
Matches 97; Conservative 74; Mismatches 122; Indels 51; Gaps 18;

```

Qy      1 DDLBOERRAKEKLOEQOSDLEODRLAKKLOEQO--SDLEOERRAKEKLOEQO--SDLEOER  58
Db      40 TESSKQSSKRTTEBELKKKQOEBELKKKQOEBELKKKQOEBELKKKQOEBELKKKQOEB  99
Qy      59 RAKEKLOEQO--SDLEOERRAKEKLOEQO--SDLEODRLAKKLOEQOSDLEOERRAKEK  115
Db      100 ELKKKQOEBELKKKQOEBELKKKQOEBELKKKQOEBELKKKQOEBELKKKQOEBELKKKQ  159
Qy      116 QEOO--SDLEOERRAKEKLOEQO--SDLEOERRAKEKLOEQO--SDLEOERRAKEKLOEQO--S  171
Db      160 QOEBELKKKQOEBELKKKQOEBELKKKQOEBELKKKQOEBELKKKQOEBELKKKQOEBELK  219
Qy      172 DLEOERRAKEKLOEQOSDLEOERRAKEKLOEQO-----RDLEOR-----KADTKNUL  218
Db      220 KKOQEBELKKKQOEBELKKKQOEBEKEKELRKQOLNLTLSNDLKQOIBASVNFKEKYVXSM  279
Qy      219 ERKKEHGDIALDYLRLI-----PAIELPSEN-ENGY-YIPQSSL--PODNRGMSR  268
Db      280 EKE-----PEDHGMTSFRLNMGPTBEDISDTERISIRYRRTTYIVLSPLD----PH  328
Qy      269 DSKEL-SIIETKNRESITTN-----VEGRDIHKGHLEKXD  304
Db      329 ELKEFPANIIODINKLASVASIFNSFSATGAIGALDVIDSLYERKD  372

```

RESULT 10
E84565
hypothetical protein. At2g18540 (imported) - Arabidopsis thaliana

C:/Species: *Arabidopsis thaliana* (mouse-ear cress)
C:/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:/Accession: E84565
R./Lin. X.: Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Mofett, K.S.; Cronin, L.A.; Shen, M.; Vannken, S.E.; Unayam, L.; Tallon, L.;
euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A./Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A./Reference number: A84420; MUID:20083487; PMID:10617129
A./Accession: E84565
A./Status: preliminary
A./Molecule type: DNA
A./Residues: 1-699 <STO>
A./Cross-references: UNIPROT:Q9ZU69; UNIPARC:UPI00000A5E25; GB:AE002093; NID:g4218005; P
C./Genetics:
A./Gene: At2g18540
A./Map position: 2

Query Match	17.4%	Score 272:	DB 2:	Length 699;
Best Local Similarity	28.9%	Pred. No. 5.8e-06;		
Matches 74; Conservative	69;	Mismatches 87;	Indels 26;	Gaps 9;

Qy 4 QOERRAKKLOEOQSDLEODRL--AKKLOEQSDLEOERRAKKLOEQSDLEOERRA 60
 Db 426 ERKRREBEIERRRKREBEAKKREBAKREBEBAKREBEETERKREBEERAKKREBERK 485
 Qy 61 KEKLOEQSDLEOERRAKKLOEQSDLEODRLAKKLOEQSDLEOERRAKKLO-Q-EOQ 119
 Db 486 RE--EERAKRREBEKKREBEBAQARKREER--EKEEWAKKREBERORKREBEVERK 540
 Qy 120 SDLEERRAKKLOEQSDLEOERRAKKLOEQSDLEOERRAKK-----KLOEQSDL 173
 Db 541 RREDOERRKRE--EERAKREBERKREBEEMAKR--EOEROKREBEVERKIREBOERK 594
 Qy 174 EOERRAKKLOEQSDLEOERRAKKLOEQSDLEORAKDTKONLERKKEHGLIADLY 233
 Db 595 REEWAKKREBOERQCK-EREWERKCKREBEAKKREEWAKKIREERQCKEREDVERK--- 650
 Qy 224 GRLEIPALEPSENER 249
 Db 651 -RREBAKREBERCKR 665

RESULT 11

```

protein.F23P12.8 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: G88484
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99063613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_el
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: G88484
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-987 <STO>
A:Cross-references: UNIPROT:P46504; UNIPARC:UPI0000060FC9; GB:chr_III; PIDN:AAA20610.1;
A:Note: exon 5 similar to trichohyalin and tropomyosin
C:Genetics:
A:Gene: F23P12.8
A:Map position: 3

```

Query Match	17.4%	Score 271.5	DB 2	Length 887
Best Local Similarity	25.0%	Pred. No. 7.5e-06		
Matches 83, Conservative	93	Mismatches 105	Indels 51	Gaps 13

```

Oy      6 RRRRAKEKLOEQOSDLEODRLAKEKLOEQOSDLEOERRAKEKLOEQOSDLE-----EOE 57
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      310 EROOQEKER---MEORLPOEK-BEKARELERRKLEESBTARQAEIDROATTIYABOE 364
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      58 RRAKEKLOE-QOSDLEOERRAKEKLOEQOSDLEODRLAK-EXTLOEQOSDLEOER--RAK 112
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      365 RMAHERNELERIRLIEKKRENERROEIMAEISKILBELRLQ-----LEORQNNERAR 419
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      113 EKLQEQOSDLEOERRAKEKLOEQOSDLEOERRAKEKLOEQOSDLEOERRAKEKLOEQOSD 172
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      420 QELEARKYKLOEBRQKIQOKKWEWOIQOEBARQOULVLEEB-RARELERVQOE 478
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      173 LEORRAKEKLOEQOSDLEOERRAKEKLOEQORD-----LEORKADTKNLERKKEH 224
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      479 LERHOQ-MEILROQOEBQKKKLEEDROQRORQOEAEELNRMIIEKEMKENNQKITEKQK 537
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      225 GDILAEDLYGELIPATELPSENERGYIIPHQSSLPODNRGNSRKSISITIEKTN-RES 283
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      538 RKMLEKEN-----EDROQATYEEEBERRIAEEER---RKQIEIEERRRIOQ 580
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      284 ITTNVEGRDIIHGKLEKKDGSIKPEQEKER 315
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      581 IMATBERSRLDAMEREREMLRQITESEKQK 612
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12

IGA_receptor - Streptococcus pyogenes
 C:Species: Streptococcus pyogenes
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
 C:Accession: S37046
 R:Heden, L.; Lindahl, G.
 submitted to the EMBL Data Library, May 1993
 A:Description: Conserved and variable regions in protein Arp, the IGA receptor of Streptococcus pyogenes
 A:Reference number: S37046
 A:Accession: S37046
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-402 <HED>
 A:Cross-references: UNIPROT:Q54876; UNIPARC:UPI00000BC74D; EMBL:Z22751; NID:G397628; PIR:G39762
 C:Superfamily: M5 protein
 i:Keywords: Immunoglobulin receptor

Query Match	17.2%	Score 269;	DB 2;	Length 402;
Best Local Similarity	27.9%	Pred. No. 5e-06;		
Match 93; Conservative	63;	Mismatches 105;	Indels 72;	Gaps 12;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:11:08 ; Search time 99.9815 Seconds
(without alignments)
2229.881 Million cell updates/sec

Title: US-09-837-344-31
Perfect score: 1563
Sequence: 1 SDLEQERRAKKLEQEQSD.....GHLEKKDSIKPEQKEDKS 316

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1549	99.1	493	Q25886_PLAFA	Q25886 plasmodium
2	1494	95.6	1909	Q25893_PLAFA	Q25893 plasmodium
3	1453.5	93.0	1596	Q81J44_PLAF7	Q81J44 plasmodium
4	771.5	49.4	318	Q25887_PLAFA	Q25887 plasmodium
5	554	35.4	280	Q25844_PLAFA	Q25844 plasmodium
6	554	35.4	280	Q25849_PLAFA	Q25849 plasmodium
7	554	35.4	280	Q25850_PLAFA	Q25850 plasmodium
8	553	35.4	280	Q25843_PLAFA	Q25843 plasmodium
9	553	35.4	280	Q25846_PLAFA	Q25846 plasmodium
10	553	35.4	280	Q25847_PLAFA	Q25847 plasmodium
11	553	35.4	280	Q25851_PLAFA	Q25851 plasmodium
12	550	35.2	280	Q25848_PLAFA	Q25848 plasmodium
13	550	35.2	280	Q25852_PLAFA	Q25852 plasmodium
14	549	35.1	280	Q25845_PLAFA	Q25845 plasmodium
15	549	35.1	280	Q25853_PLAFA	Q25853 plasmodium
16	549	35.1	280	Q25854_PLAFA	Q25854 plasmodium
17	549	35.1	280	Q25855_PLAFA	Q25855 plasmodium
18	549	35.1	280	Q25888_PLAFA	Q25888 plasmodium
19	549	35.1	280	Q25889_PLAFA	Q25889 plasmodium
20	549	35.1	280	Q25900_PLAFA	Q25900 plasmodium
21	549	35.0	280	Q27243_PLAFA	Q27243 plasmodium
22	469	30.0	264	Q9GTX5_PLAFA	Q9GTX5 plasmodium
23	416	26.6	978	Q7RAK4_PLAYO	Q7RAK4 plasmodium
24	387.5	24.8	2205	Q54C75_DICDI	Q54C75 dictyosteli
25	379	24.2	1108	Q9ND10_BABBI	Q9ND10 babesia big
26	378.5	24.2	1207	Q4UHB4_THEAN	Q4UHB4 theileria a
27	372.5	23.8	905	Q54XN9_DICDI	Q54XN9 dictyosteli
28	372.5	23.8	1777	Q54FB8_DICDI	Q54FB8 dictyosteli
29	365.5	23.4	3427	Q6YA79_PLARE	Q6YA79 plasmodium
30	364.5	23.3	826	Q9VPS3_DROME	Q9VPS3 drosophila
31	360	23.0	1575	Q519G4_ENTHI	Q519G4 entamoeba h

32	359	23.0	1082	Q54UG8_DICDI	Q54UG8 dictyosteli
33	358	22.9	762	Q50RA3_ENTHI	Q50RA3 entamoeba h
34	355	22.7	1003	Q91IX9_HNVB	Q91IX9 human herpe
35	354.5	22.7	1129	Q9QR71_HNVB	Q9QR71 human herpe
36	353.5	22.6	1036	Q9DUM3_HNVB	Q9DUM3 human herpe
37	352	22.5	1080	Q559H7_DICDI	Q559H7 dictyosteli
38	350.5	22.4	899	Q7KWT7_DICDI	Q7KWT7 dictyosteli
39	349.5	22.4	1069	Q40947_HNVB	Q40947 human herpe
40	349	22.3	1059	Q51AX0_ENTHI	Q51AX0 entamoeba h
41	342	21.9	976	Q9DUN0_HNVB	Q9DUN0 human herpe
42	342	21.9	1119	Q54TT8_DICDI	Q54TT8 dictyosteli
43	334	21.3	1061	Q54MJ1_DICDI	Q54MJ1 dictyosteli
44	332.5	21.3	1418	Q86HT4_DICDI	Q86HT4 dictyosteli
45	332.5	21.3	1466	Q554X7_DICDI	Q554X7 dictyosteli

ALIGNMENTS

RESULT 1					
Q25886_PLAFA	PRELIMINARY;	PRT;	493 AA.		
ID Q25886_PLAFA					
AC Q25886;					
DT 01-NOV-1996 (TREMBLrel. 01, Created)					
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)					
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)					
DE Liver stage antigen-1 (Fragment).					
GN Name=LSA-1;					
OS Plasmodium.					
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.					
OX NCBI_TaxID=5633;					
RN [1]					
RP NUCLEOTIDE SEQUENCE.					
RX MEDLINE=9426724; Pubmed=7515922;					
RA Fidock D.A., Gras-Masse H., Lepers J., Brahimi K., Benmohamed L.,					
RA Meis J.F., Langsley G., Rousillon C., Londono A., Raharimalala L.,					
RA Meis J.F., Langsley G., Rousillon C., Tartar A., Drulhe P.,					
RT "Plasmodium falciparum liver stage antigen-1 is well conserved and					
RT contains potent B and T cell determinants.";					
RL J. Immunol. 153:190-204 (1994).					
DR EMBL; Z30320; CAA82975.1; -; Genomic_DNA.					
FT NON_TER					
SO SEQUENCE	493 AA;	59062 MW;	FE252CC44F46996E CRC64;		
Query Match					
Best Local Similarity 99.1%; Score 1549; DB 2; Length 493;					
Matches 314; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
QY	1	SDLEQERRAKKLEQEQSDLEQDRLAKEKLEQEQSDLEQERRAKKLEQEQSDLEQERRA	60		
DB	5	SDLEQERRAKKLEQEQSDLEQDRLAKEKLEQEQSDLEQERRAKKLEQEQSDLEQERRA	64		
QY	61	KEKLEQEQSDLEQERRAKKLEQEQSDLEQDRLAKEKLEQEQSDLEQERRAKKLEQEQSDLEQERRA	120		
DB	65	KEKLEQEQSDLEQERRAKKLEQEQSDLEQDRLAKEKLEQEQSDLEQERRAKKLEQEQSDLEQERRA	124		
QY	121	DLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRA	180		
DB	125	DLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRA	184		
QY	181	EKLQEQSDLEQERRAKKLEQEQSDLEQDRADTKNLERKEHGDILAEDLYGRLEIPA	240		
DB	185	EKLQEQSDLEQERRAKKLEQEQSDLEQDRADTKNLERKEHGDILAEDLYGRLEIPA	244		
QY	241	IEPSENERGYIYIHQSSLPQDNNGNSRDSKEISIEKTNRRESITTVNVEGRDIIHKHLE	300		

```

Db      245 IELPSENERGYIIPHOSSLPODNRGSRDSKEISITTEKTNRESITTNVEGRDIIKHGHL 304
Qy      301 EKKDGSIKPEQKEDKS 316
Db      305 EKKDGSIKPEQKEDKS 320

RESULT 2
Q25893 PLAF7
ID Q25893 PLAF7 PRELIMINARY; PRT; 1909 AA.
AC Q25893;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Liver stage antigen.
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NF54;
RX MEDLINE=92107224; PubMed=1840628; DOI=10.1016/0166-6851(91)90117-O;
RA Zhu J., Hollingdale M.R.;
RT "Structure of Plasmodium falciparum liver stage antigen-1."
RL Moll. Biochem. Parasitol. 48:223-226 (1991).
DR EMBL; X56203; CA39663.1; -; Genomic_DNA.
DR PIR; S24597; A45592.
DR HSSP; O15813; 1D7M.
SQ SEQUENCE 1909 AA; 230153 MW; 8B861A524C797217 CRC64;

Query Match      95.6%; Score 1494; DB 2; Length 1909;
Best Local Similarity 95.9%; Pred. No. 2,4e-50;
Matches 303; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy      1 SDLEORRAKEKLEQOQSDLEODRLAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 60
Db      1421 SDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 1480
Qy      61 KEKLEQOQSDLEORRAKEKLEQOQSDLEODRLAKEKLEQOQSDLEORRAKEKLEQOQ 120
Db      1481 KEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQ 1540
Qy      121 DLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 180
Db      1541 DLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 1600
Qy      181 EKLQEQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLE 240
Db      1601 EKLQEQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLE 1660
Qy      241 IELPSENERGYIIPHOSSLPODNRGSRDSKEISITTEKTNRESITTNVEGRDIIKHGHL 300
Db      1661 IELPSENERGYIIPHOSSLPODNRGSRDSKEISITTEKTNRESITTNVEGRDIIKHGHL 1720
Qy      301 EKKDGSIKPEQKEDKS 316
Db      1721 EKKDGSIKPEQKEDKS 1736

RESULT 3
Q25893 PLAF7
ID Q25893 PLAF7 PRELIMINARY; PRT; 1596 AA.
AC Q25893;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Liver stage antigen, putative.
GN ORFNames=PF10_0356;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=36329;

```

```

RN      [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angioli S.,
RA Pereira M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carnocci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Frazer C.W., Barrell B.G.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511 (2002).
DR EMBL; AF014834; AAN35553.1; -; Genomic DNA.
SQ SEQUENCE 1596 AA; 191691 MW; C0B3A2CC0F5D60AC CRC64;

Query Match      93.0%; Score 1453.5; DB 2; Length 1596;
Best Local Similarity 90.1%; Pred. No. 7.7e-49;
Matches 300; Conservative 5; Mismatches 11; Indels 17; Gaps 1;

Qy      1 SDLEORRAKEKLEQOQSDLEODRLAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 60
Db      1091 SDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 1150
Qy      61 KEKLEQOQSDLEORRAKEKLEQOQSDLEODRLAKEKLEQOQSDLEORRAKEKLEQOQ 120
Db      1151 KEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQ 1210
Qy      121 DLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 163
Db      1211 DLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 1270
Qy      164 EKLQEQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLE 223
Db      1271 EKLQEQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLE 1330
Qy      224 HGDVLAEDLYGRLPALELPSENERGYIIPHOSSLPODNRGSRDSKEISITTEKTNRES 283
Db      1331 HGDVLAEDLYGRLPALELPSENERGYIIPHOSSLPODNRGSRDSKEISITTEKTNRES 1390
Qy      284 ITTNVEGRDIIKHGLEEKDGSIKPEQKEDKS 316
Db      1391 ITTNVEGRDIIKHGLEEKDGSIKPEQKEDKS 1423

RESULT 4
Q25897 PLAF7
ID Q25897 PLAF7 PRELIMINARY; PRT; 318 AA.
AC Q25897;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Liver stage antigen-1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94267224; PubMed=7515922;
RA Fidock D.A., Grais-Rasse H., Lepeyre J., Brahimi K., Benmohamed L.,
RA Wellouk S., Guerin-Marchand C., Londono A., Raharimalala L.,
RA Meis J.F., Langsley G., Roussillon C., Tartar A., Druilhe P.;
RT "Plasmodium falciparum liver stage antigen-1 is well conserved and
RT contains potent B and T cell determinants."
RL J. Immunol. 153:190-204 (1994).
DR EMBL; Z30319; CA482974.1; -; Genomic_DNA.
DR HSSP; O15813; 1D7M.
FT NON TER 318
SQ SEQUENCE 318 AA; 38098 MW; 989F32A70605DFB2 CRC64;

```

Query Match 49.4%; Score 771.5; DB 2; Length 318;
Best Local Similarity 86.2%; Pred. No. 5.3e-23;
Matches 163; Conservative 12; Mismatches 13; Indels 1; Gaps 1;

2 DLEQERRAKKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKKLEQOQSDLEQERRAK 61
131 DDDKKKXIKODENRQDLE- EKAKKLEKQOQSDSDEQERRAKKLEQOQSDLEQERLAK 189
62 EKLQEQSDLEQERRAKKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKKLEQOQSD 121
190 EKLQEQSDLEQERRAKKLEQOQSDLEQERLAKKLEQOQSDLEQERRAKKLEQOQSD 249
122 LEQERRAKKLEQOQSDLEQERRAKKLEQOQSDLEQERRAKKLEQOQSDLEQERRAK 181
250 LEQERRAKKLEQOQSDLEQERRAKKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAK 309
182 KLEQOQSDL 190
310 RLEQOQSDL 318

RESULT 5
Q25844_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25844; ID Q25844_PLAFA PRELIMINARY; PRT; 280 AA.
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RT Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40885; AAB59231.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32940 MW; E9708E3CFPAE9CF CRC64;

Query Match 35.4%; Score 554; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

210 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 269
1 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 60

QY 270 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 316
61 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 107

Db 270 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 316
61 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 6
Q25849_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25849; ID Q25849_PLAFA PRELIMINARY; PRT; 280 AA.
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RT Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40890; AAC41600.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32966 MW; E96D255154DEA9CF CRC64;

Query Match 35.4%; Score 554; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

210 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 269
1 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 60

QY 270 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 316
61 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 107

Db 270 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 316
61 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 7
Q25850_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25850; ID Q25850_PLAFA PRELIMINARY; PRT; 280 AA.
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RT Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40891; AAC41601.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32989 MW; E96F812CFPABED8 CRC64;

Query Match 35.4%; Score 554; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

210 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 269
1 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 60

QY 270 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 316
61 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 107

Db 270 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 316
61 SKEISIIKTNRESITTNVEGRDIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 8
Q25843_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25843; ID Q25843_PLAFA PRELIMINARY; PRT; 280 AA.
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

```

OX  NCBI_TaxID=5833;
RN  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RT  "Sequence variations in the non-repetitive regions of the liver stage-
RT  specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT  isolates."
RL  Mol. Biochem. Parasitol. 71:291-294(1995).
DR  EMBL, L40884; AAB59230.1; -; Genomic_DNA.
FT  NON TER
SQ  SEQUENCE 280 AA; 32926 MW; 467080932FABAAD33 CRC64;

Query Match      35.4%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.3e-14;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  210 RKADTKKLNLERKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSLPODNRGNSRD 269
    |||||
DB  1 RKADTKKLNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSLPODNRGNSRD 60

OY  270 SKEISIIKTNRESITTNVEGRDIIKHGLEKKDGSIKPEQKEDKS 316
    |||||
DB  61 SKEISIIKTNRESITTNVEGRDIIKHGLEKKDGSIKPEQKEDKS 107

RESULT 9
O25846 PLAFa
ID  O25846 PLAFa PRELIMINARY; PRT; 280 AA.
AC  O25846;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  Liver stage-specific antigen 1 (Fragment).
GN  Name=LSA-1;
OS  Plasmodium falciparum.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5833;
RN  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RT  "Sequence variations in the non-repetitive regions of the liver stage-
RT  specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT  isolates."
RL  Mol. Biochem. Parasitol. 71:291-294(1995).
DR  EMBL, L40887; AAC41597.1; -; Genomic_DNA.
FT  NON TER
SQ  SEQUENCE 280 AA; 32882 MW; 46763641C81AFC33 CRC64;

Query Match      35.4%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.3e-14;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  210 RKADTKKLNLERKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSLPODNRGNSRD 269
    |||||
DB  1 RKADTKKLNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSLPODNRGNSRD 60

OY  270 SKEISIIKTNRESITTNVEGRDIIKHGLEKKDGSIKPEQKEDKS 316
    |||||
DB  61 SKEISIIKTNRESITTNVEGRDIIKHGLEKKDGSIKPEQKEDKS 107

RESULT 10
O25847 PLAFa
ID  O25847 PLAFa PRELIMINARY; PRT; 280 AA.
AC  O25847;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  Liver stage-specific antigen 1 (Fragment).

```

```

GN  Name=LSA-1;
OS  Plasmodium falciparum.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5833;
RN  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RT  "Sequence variations in the non-repetitive regions of the liver stage-
RT  specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT  isolates."
RL  Mol. Biochem. Parasitol. 71:291-294(1995).
DR  EMBL, L40888; AAC41598.1; -; Genomic_DNA.
FT  NON TER
SQ  SEQUENCE 280 AA; 32944 MW; 467085E23BABBC27 CRC64;

Query Match      35.4%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.3e-14;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  210 RKADTKKLNLERKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSLPODNRGNSRD 269
    |||||
DB  1 RKADTKKLNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSLPODNRGNSRD 60

OY  270 SKEISIIKTNRESITTNVEGRDIIKHGLEKKDGSIKPEQKEDKS 316
    |||||
DB  61 SKEISIIKTNRESITTNVEGRDIIKHGLEKKDGSIKPEQKEDKS 107

RESULT 11
O25851 PLAFa
ID  O25851 PLAFa PRELIMINARY; PRT; 280 AA.
AC  O25851;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  Liver stage-specific antigen 1 (Fragment).
GN  Name=LSA-1;
OS  Plasmodium falciparum.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5833;
RN  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RT  "Sequence variations in the non-repetitive regions of the liver stage-
RT  specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT  isolates."
RL  Mol. Biochem. Parasitol. 71:291-294(1995).
DR  EMBL, L40893; AAC41603.1; -; Genomic_DNA.
FT  NON TER
SQ  SEQUENCE 280 AA; 32927 MW; 487065FDCP40ADD3 CRC64;

Query Match      35.4%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.3e-14;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  210 RKADTKKLNLERKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSLPODNRGNSRD 269
    |||||
DB  1 RKADTKKLNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSLPODNRGNSRD 60

OY  270 SKEISIIKTNRESITTNVEGRDIIKHGLEKKDGSIKPEQKEDKS 316
    |||||
DB  61 SKEISIIKTNRESITTNVEGRDIIKHGLEKKDGSIKPEQKEDKS 107

RESULT 12
O25848 PLAFa
ID  O25848 PLAFa PRELIMINARY; PRT; 280 AA.
AC  O25848;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)

```

DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, last annotation update)
DE Liver stage-specific antigen 1 (fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL, L40889; AAC41599.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32939 MW; E77080365FAAD61 CRC64;
Query Match 35.2%; Score 550; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.8e-14;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 210 RKADTKNLRKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 269
DB 1 RKADTKNLRKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 60
QY 270 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107
RESULT 13
ID Q25852_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25852;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, last annotation update)
DE Liver stage-specific antigen 1 (fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL, L40834; AAB59233.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32927 MW; E6648F85FBA245 CRC64;
Query Match 35.2%; Score 550; DB 2; Length 280;
Best Local Similarity 98.1%; Pred. No. 1.8e-14;
Matches 105; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 210 RKADTKNLRKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 269
DB 1 RKADTKNLRKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 60
QY 270 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107
RESULT 14
ID Q25845_PLAFA PRELIMINARY; PRT; 280 AA.

ID Q25845_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25845;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, last annotation update)
DE Liver stage-specific antigen 1 (fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL, L40886; AAB59232.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32908 MW; FDEB929F34DCAT99 CRC64;
Query Match 35.1%; Score 549; DB 2; Length 280;
Best Local Similarity 98.1%; Pred. No. 1.9e-14;
Matches 105; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 210 RKADTKNLRKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 269
DB 1 RKADTKNLRKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 60
QY 270 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107
RESULT 15
ID Q25853_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25853;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, last annotation update)
DE Liver stage-specific antigen 1 (fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL, L40835; AAB59234.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 46751C45F4DCBD33 CRC64;
Query Match 35.1%; Score 549; DB 2; Length 280;
Best Local Similarity 98.1%; Pred. No. 1.9e-14;
Matches 105; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 210 RKADTKNLRKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 269
DB 1 RKADTKNLRKKEHGDILAEADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 60
QY 270 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107

Search completed: December 7, 2005, 07:20:44
Job time : 99.9815 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 06:57:58 ; Search time 102.171 Seconds
(without alignments)
1358.936 Million cell updates/sec

Title: US-09-837-344-31

Perfect score: 1563
Sequence: 1 SDLEQERRAKEXKQEQSD.....GHLEKKDSIKPEQKEDKS 316

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq 21: *
1: geneseqp19908: *
2: geneseqp19908: *
3: geneseqp20008: *
4: geneseqp20018: *
5: geneseqp20028: *
6: geneseqp20038: *
7: geneseqp20038: *
8: geneseqp20048: *
9: geneseqp20058: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1549	99.1	316	2	AAR26941 P.falcipa
2	1549	99.1	493	2	AAR26944 P.falcipa
3	942.5	60.3	462	2	AAR05766 Portion o
4	771.5	49.4	318	2	AAR26943 P.falcipa
5	739	47.3	456	8	ADO21942 LSA-NRC(H
6	739	47.3	457	8	ADO21920 LSA-NRC(H
7	603	38.6	117	2	AAR26937 P.falcipa
8	331	21.2	207	3	AAB44672 Gene 33 h
9	329.5	21.1	1439	8	ADRI12189 Trichocho
10	329.5	21.1	1439	8	ADU77050 Murine ha
11	329.5	21.1	1439	9	ADM00404 Epithelia
12	327.5	21.0	1162	3	AAV96255 Kaposi's
13	327.5	21.0	1162	3	AAV96255 Kaposi's
14	327.5	21.0	1162	4	AAV96255 Kaposi's
15	327.5	21.0	1162	5	ABR05621 Kaposi's
16	327.5	21.0	1162	8	ADJ65096 HHV8 late
17	327.5	21.0	1162	9	ADJ65096 HHV8 late
18	326	20.9	611	2	AAV29039 T. gondii
19	326	20.9	611	4	AAU25510 T. gondii
20	326	20.9	611	7	ADG17131 T. gondii
21	321	20.5	360	2	AAW03627 Human fol
22	321	20.5	412	2	AAW03627 Human thy
23	321	20.5	554	4	ABR59454 Drosophi
24	317.5	20.3	1898	2	AAV30795 A human t

25	317.5	20.3	1898	7	ADD48869	Add48869 Human pro
26	315.5	20.2	1584	8	ABM80133	Abm80133 Tumour-as
27	312	20.0	345	4	ABU53267	Abu53267 Human tes
28	311	19.9	621	7	ADM05724	Adm05724 Human pro
29	309	19.8	839	8	ADU02517	Adu02517 Novel hum
30	306	19.6	2274	4	ABR58657	Abbr58657 Drosophi
31	302.5	19.3	562	2	AAR70491	Aar70491 Leucocyto
32	301	19.3	1180	7	ADP06151	Adp06151 Bacterial
33	296.5	19.0	2980	8	ADP29884	Adp29884 Human sec
34	289	18.5	407	8	ADP49328	Adp49328 S pyrogen
35	289	18.5	923	4	ABG08600	Abg08600 Novel hum
36	288	18.4	650	7	ADB65482	Adb65482 Human pro
37	288	18.4	949	7	ABG15508	Abg15508 Novel hum
38	285	18.2	1940	8	ADP29883	Adp29883 Human sec
39	284	18.2	67	1	AAPE81355	Aape81355 Sequence
40	284	18.2	1715	4	ABR58089	Abbr58089 Drosophi
41	282	18.0	466	7	ADM05659	Adm05659 Human pro
42	281	18.0	658	4	ABR65632	Abbr65632 Drosophi
43	281	18.0	800	4	ABR71459	Abbr71459 Drosophi
44	280.5	17.9	345	8	ABO60621	Abbo60621 Human gen
45	278	17.8	280	7	ADB64909	Adb64909 Human pro

ALIGNMENTS

RESULT 1	
AAR26941	
ID	AAR26941 standard; protein; 316 AA.
XX	
AC	AAR26941;
XX	
DT	25-MAR-2003 (revised)
DT	08-FEB-1993 (first entry)
XX	
DE	P.falciparum LSA-R-NR protein.
XX	
KW	Malaria; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;
KW	paludism; liver stage-specific antigen.
XX	
OS	Plasmodium falciparum.
XX	
FN	Key
FN	Region
FT	Location/Qualifiers
FT	1..209
FT	/label=repeat_region
FT	/note="contains 12 x 17mer repeats"
FT	210..316
FT	/label=non-repeat_region
XX	
PN	W09213884-A1.
XX	
PD	20-AUG-1992.
XX	
PF	05-FEB-1992; 92WC-FR000104.
XX	
PR	05-FEB-1991; 91FR-00001286.
XX	
PA	(INSP) INST PASTEUR.
XX	
PI	Guerinmarchand C, Drulhe P;
XX	
DR	WPI, 1992-299985/36.
DR	N-PSDB; AAQ28115.
XX	
PT	Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT	vaccination against, treatment of and diagnosis of malaria.
XX	
PS	Disclosure; Fig 1; 81pp; French.
XX	
CC	A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
CC	gt11 was used to transform E.coli. The expression library was screened
CC	with human antisera against antigens of all stages of P. falciparum. The
CC	library was rescreened with antibodies affinity-purified on a clone

CC which was able to recognise antibodies specific to the hepatic phase.
CC About 40 clones were detected which produced a characteristic USA
CC epitope. The clone with the largest insert (950 bases) encoded USA-R-NR
CC containing a 12-repeat region followed by a non-repeat region. Preferred
CC antigenic polypeptides of the invention are derived from the amino acid
CC sequence of USA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 316 AA;

Query Match 99.1%; Score 1549; DB 2; Length 316;
Best Local Similarity 99.4%; Pred. No. 1.3e-109;
Matches 314; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOSDLEQERRA 60
DB 1 SDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOSDLEQERRA 60
QY 61 KEKLEQOSDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOS 120
DB 61 KEKLEQOSDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOS 120
QY 121 DLEERRAKKELQEQOSDLEERRAKKELQEQOSDLEERRAKKELQEQOSDLEERRAK 180
DB 121 DLEERRAKKELQEQOSDLEERRAKKELQEQOSDLEERRAKKELQEQOSDLEERRAK 180
QY 181 EKLQEQOSDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOS 240
DB 181 EKLQEQOSDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOS 240
QY 241 IELPSENERGYIIPHOSSLPODNRGNSRDSKEISIIIEKTNRESITTNVGRDDIHKGHLE 300
DB 241 IELPSENERGYIIPHOSSLPODNRGNSRDSKEISIIIEKTNRESITTNVGRDDIHKGHLE 300
QY 301 EKKDGSIRPEQKEDKS 316
DB 301 EKKDGSIRPEQKEDKS 316

RESULT 2

AAR26944
ID AAR26944 standard; protein; 493 AA.

XX AAR26944;

DT 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)

DE P.falciparum USA gene C-terminal region.

KW Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
KW paludism; liver stage-specific antigen.

OS Plasmodium falciparum.

Key Location/Qualifiers

FT Region 13..213
FT /label= repetitive_region
FT Region 214..493
FT /label= non-repetitive_region

PN WO9213884-A1.

XX 20-AUG-1992.

PF 05-FEB-1992; 92MO-FR000104.

PR 05-FEB-1991; 91FR-00001286.

XX (INSP) INST PASTEUR.

PI Guerinmarchand C, Drulhe P;

XX WPI, 1992-299985/36.

DR N-PSDB; AAQ28119.

XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.

XX Claim 2; Fig 8-10; 81pp; French.

CC The 3' part of the P.falciparum liver-stage specific antigen (USA) gene
CC codes for a polypeptide sequence which carries a T cell epitope
CC characteristic of a protein produced in hepatocytes infected with
CC P.falciparum. The polypeptide can be used in the preparation of vaccines
CC against malaria. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 493 AA;

Query Match 99.1%; Score 1549; DB 2; Length 493;
Best Local Similarity 99.4%; Pred. No. 2e-109;
Matches 314; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOSDLEQERRA 60
DB 5 SDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOSDLEQERRA 64
QY 61 KEKLEQOSDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOS 120
DB 65 KEKLEQOSDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOS 124
QY 121 DLEERRAKKELQEQOSDLEERRAKKELQEQOSDLEERRAKKELQEQOSDLEERRAK 180
DB 125 DLEERRAKKELQEQOSDLEERRAKKELQEQOSDLEERRAKKELQEQOSDLEERRAK 184
QY 181 EKLQEQOSDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOS 240
DB 185 EKLQEQOSDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOS 244
QY 241 IELPSENERGYIIPHOSSLPODNRGNSRDSKEISIIIEKTNRESITTNVGRDDIHKGHLE 300
DB 245 IELPSENERGYIIPHOSSLPODNRGNSRDSKEISIIIEKTNRESITTNVGRDDIHKGHLE 304
QY 301 EKKDGSIRPEQKEDKS 316
DB 305 EKKDGSIRPEQKEDKS 320

RESULT 3

AAR05766
ID AAR05766 standard; protein; 462 AA.

XX AAR05766;

DT 25-MAR-2003 (revised)
DT 05-NOV-1990 (first entry)

DE Portion of peptide antigen to malarial sporozoite.

KW Malaria; sporozoite; vaccine; exoerythrocytic parasites; tetanus toxoid.

OS Plasmodium falciparum.

PN WO9006130-A.

PD 14-JUN-1990.

PF 30-NOV-1988; 88US-00278234.

PR 30-NOV-1988; 88US-00278234.

XX 12-APR-1989; 89US-00337204.

XX (BIOM-) BIOMEDICAL RES INST.

PI Hollingdal MR;

XX WPI, 1990-209624/27.


```

PT Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
XX (LSA-1) epitope.
XX
PS Claim 4; SEQ ID NO 26; 90pp; English.
XX
CC The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct protein of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to a C-terminal His6 tag.
XX
SQ Sequence 456 AA;
Query Match 47.3%; Score 739; DB 8; Length 456;
Best Local Similarity 55.1%; Pred. No. 5,6e-48;
Matches 163; Conservative 32; Mismatches 71; Indels 30; Gaps 5;
QY 21 EODRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLE
DB 6 EKDEIKSNLKGSSN-----SKRINIEKHKKHVSINSHSYETKNNENK 52
QY 81 LQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLE
DB 53 FPDKCKELTMSNV-KNVSQTNFKSLRLNIGVSENIFLKENKLNKKGKLIENHINDDD-- 109
QY 141 QERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLE
DB 110 KKKYIKGQDENRQEDLE-----EKAAEQQSDLEQERRAKKLE-----QERRAKKLE 155
QY 201 QEQGRDLEQKADTYKKNLERKKEHGDILAEDLYGRLEIPIALEPSENERGYIIPHOSSLP 260
DB 156 QEQGRDLEQKADTYKKNLERKKEHGDVLAEDLYGRLEIPIALEPSENERGYIIPHOSSLP 215
QY 261 QDNRGNRSRDSKEISIIETKNRESITTNVGRDIIKHGLEKKDGSIKPEQKEDKS 316
DB 216 QDNRGNRSRDSKEISIIETKNRESITTNVGRDIIKHGLEKKDGSIKPEQKEDKS 271
RESULT 6
ADO21920
ID ADO21920 standard; protein; 457 AA.
XX
AC ADO21920;
XX
DT 12-AUG-2004 (first entry)
XX
DE LSA-NRC(H)Mut construct protein derived from Malaria parasite LSA-1.
XX
KM LSA-NRC, liver stage antigen; LSA-1; T-cell; B-cell epitope;
KM parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
KM malaria parasite P. falciparum; LSA-NRC(H)Mut construct; mutant; mutein.
XX
OS Plasmodium falciparum.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 433
FT /note="Residue inserted relative to wild-type P.
FT falciparum T5 epitope"
XX
PN WO2004044167-A2.
XX
PD 27-MAY-2004.
XX
PF 12-NOV-2003; 2003WO-US036011.
XX
PR 12-NOV-2002; 2002US-0425719P.

```

```

XX
PA (REED-) REED ARMY INST RES WALTER.
XX
FI Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Barbosa A;
XX
DR WPI; 2004-420309/39.
DR N-PSDB; ADO21919.
XX
PT Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
XX (LSA-1) epitope.
XX
PS Claim 5; SEQ ID NO 4; 90pp; English.
XX
CC The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct protein of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to an amino acid insertion
CC within the T5 P. falciparum epitope and a C-terminal His6 tag.
XX
SQ Sequence 457 AA;
Query Match 47.3%; Score 739; DB 8; Length 457;
Best Local Similarity 55.1%; Pred. No. 5,6e-48;
Matches 163; Conservative 32; Mismatches 71; Indels 30; Gaps 5;
QY 21 EODRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLE
DB 6 EKDEIKSNLKGSSN-----SKRINIEKHKKHVSINSHSYETKNNENK 52
QY 81 LQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLE
DB 53 FPDKCKELTMSNV-KNVSQTNFKSLRLNIGVSENIFLKENKLNKKGKLIENHINDDD-- 109
QY 141 QERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLE
DB 110 KKKYIKGQDENRQEDLE-----EKAAEQQSDLEQERRAKKLE-----QERRAKKLE 155
QY 201 QEQGRDLEQKADTYKKNLERKKEHGDILAEDLYGRLEIPIALEPSENERGYIIPHOSSLP 260
DB 156 QEQGRDLEQKADTYKKNLERKKEHGDVLAEDLYGRLEIPIALEPSENERGYIIPHOSSLP 215
QY 261 QDNRGNRSRDSKEISIIETKNRESITTNVGRDIIKHGLEKKDGSIKPEQKEDKS 316
DB 216 QDNRGNRSRDSKEISIIETKNRESITTNVGRDIIKHGLEKKDGSIKPEQKEDKS 271
RESULT 7
AAR26937
ID AAR26937 standard; peptide; 117 AA.
XX
AC AAR26937;
XX
DT 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX
DE P.falciparum liver-stage specific antigen non-repeat region.
XX
KM Malaria; LSA-NR; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;
KM paludism.
XX
OS Plasmodium falciparum.
XX
PN WO9213884-A1.
XX
PD 20-AUG-1992.

```

```
XX 05-FEB-1992; 92MO-FR000104.
XX 05-FEB-1991; 91FR-00001286.
XX (INSP ) INST PASTEUR.
XX Guerimarchand C, Druilhe P;
XX WPI; 1992-299985/36.
XX
XX Polypeptide(6) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX vaccination against, treatment of and diagnosis of malaria.
XX
XX Claim 4; Page 55; 81pp; French.
XX
XX A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
XX CC gIII was used to transform E.coli. The expression library was screened
XX CC with human antisera against antigens of all stages of P. falciparum. The
XX CC library was rescreened with antibodies affinity-purified on a clone
XX CC which was able to recognise antibodies specific to the hepatic phase.
XX CC About 40 clones were detected which produced a characteristic LSA
XX CC epitope. The clone with the largest insert encoded LSA-R-NR containing a
XX CC 12-repeat region (LSA-R) followed by a non-repeat region (LSA-NR). The
XX CC sequence given here is a preferred polypeptide of the invention carrying
XX CC a T cell epitope typical of liver-stage P.falciparum. See AAR26919
XX CC R26935; AAR26937 and AAR26939. (Updated on 25-MAR-2003 to correct PN
XX CC field.)
XX
XX Sequence 117 AA;
SQ
Query Match 38.6%; Score 603; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 2,7e-38;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 200 LQEOQRDLQORAKDPKKNLKKKKKKGGDILAEEDLYGRLIPALPSENERGYYPHOSSL 259
DB 1 LQEOQRDLQORAKDPKKNLKKKKKKGGDILAEEDLYGRLIPALPSENERGYYPHOSSL 60
OY 260 PDNRNGNSRDSKEISIIKTNRESITTVNVEGRDIIHGKLEBKDGSIKPEQEDKS 316
DB 61 PDNRNGNSRDSKEISIIKTNRESITTVNVEGRDIIHGKLEBKDGSIKPEQEDKS 117
RESULT 8
AAB44672
ID AAB44672 standard; protein; 207 AA.
XX
XX AAB44672;
XX
XX 12-FEB-2001 (first entry)
XX
XX Gene 33 human secreted protein homologous amino acid sequence #137.
XX
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX KM noctropic; neuroprotective; antiviral; antileptergic; hepatotropic;
XX KM antidiabetic; antiinflammatory; anticancer; vulnerary; anticonvulsant;
XX KM antibacterial; antifungal; antiparasitic; cardiac; gene therapy;
XX KM food additive; preservative; chromosome identification; cancer;
XX KM immune disorder; cardiovascular disorder; neurological disease;
XX KM wound healing; infectious disease.
XX
XX Homo sapiens.
XX
XX MO200058339-A2.
XX
XX 05-OCT-2000.
XX
XX 22-MAR-2000; 2000MO-US007440.
XX
XX 26-MAR-1999; 99US-0126503P.
XX 17-DEC-1999; 99US-0172409P.
XX
```

```
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsu G;
XX
XX WPI; 2000-594637/56.
XX
XX Fifty nucleic acid molecules encoding human secreted proteins, useful in
XX PT the prevention, treatment and diagnosis of cancer, immune disorders,
XX PT cardiovascular disorders and neurological diseases.
XX
XX Disclosure; Page 399-400; 410pp; English.
XX
XX The polynucleotide sequences given in AAC79681 to AAC79730 encode the
XX CC human secreted proteins given in AAB44596 to AAB44645. AAB44646 to
XX CC AAB44693 represent human secreted polypeptide sequences and proteins
XX CC homologous to them, which are given in the exemplification of the present
XX CC invention. Human secreted proteins have activities based on the tissues
XX CC and cells the genes are expressed in. Examples of activities include:
XX CC cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
XX CC antileptergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
XX CC vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
XX CC cardiac. The polynucleotides and polypeptides are useful for preventing,
XX CC treating or ameliorating a medical condition in e.g. humans, mice,
XX CC rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
XX CC can also be used as a food additive or preservative to increase or
XX CC decrease storage capabilities. The polynucleotides are useful for
XX CC chromosome identification. They are also useful as probes for diagnosing
XX CC a disorder related to the female reproductive system, particularly breast
XX CC and/or ovarian cancer. They are also useful in the gene therapy of breast
XX CC and ovarian cancer. The nucleic acids, protein, antibodies, agonists and
XX CC antagonists from the present invention are useful in the diagnosis,
XX CC treatment and prevention of: cancer; immune disorders; cardiovascular
XX CC disorders; wound healing; neurological diseases; and infectious diseases.
XX CC AAC79672 to AAC79680 and AAB44595 represent sequences used in the
XX CC exemplification of the present invention
XX
XX Sequence 207 AA;
SQ
Query Match 21.2%; Score 331; DB 3; Length 207;
Best Local Similarity 31.7%; Pred. No. 2,2e-17;
Matches 70; Conservative 90; Mismatches 33; Indels 28; Gaps 11;
OY 4 EOBERRAKEKLOEQOSDLEQDRRAKELQEQOSDLEQERRAKEKLOEQOSDLEQERRAKE 62
DB 6 EQQOQDEQOQDEQOQD-EQQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQO 64
OY 63 KLOEQOSDLEQERRAKEKLOEQOSDLEQDRRAKELQEQOSDLEQERRAKEKLOEQOSD 122
DB 65 ELEQOQLEQEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDE 115
OY 123 EQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKE 182
DB 116 EQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQO 163
OY 183 LQEOQRDLQORAKDPKKNLKKKKKKGGDILAEEDLYGRLIPALPSENERGYYPHOSSL 223
DB 164 LQEOQRDLQORAKDPKKNLKKKKKKGGDILAEEDLYGRLIPALPSENERGYYPHOSSL 199
RESULT 9
ADR12189
ID ADR12189 standard; protein; 1439 AA.
XX
XX ADR12189;
XX
XX 21-OCT-2004 (first entry)
XX
XX Trichohyalin related protein, SEQ ID 1.
XX
XX Hair growth; murine; trichohyalin.
XX
XX Mus musculus.
XX
```

PN WO2004063750-A1.
 XX 29-JUL-2004.
 XX 07-JAN-2004; 2004WO-JP000030.
 XX 08-JAN-2003; 2003JP-00001891.
 XX (SUME) SUMITOMO ELECTRIC IND LTD.
 XX Takebe K, Hirai Y;
 XX WPI: 2004-571498/55.
 DR N-PSDB; ADRI2190.
 PT Evaluating hair-promoting activity, by using antibody that recognizes
 PT proteins (antigen) of epithelial newborn follicles.
 XX
 XX Claim 1; SEQ ID NO 1; 39pp; Japanese.
 CC The present invention relates to a method (M1) for evaluating hair growth
 CC -promoting activity, by using an antibody that recognizes an antigen
 CC (ADRI2189) of epithelial newborn follicles. Also claimed is a kit for
 CC evaluating hair growth-promoting activity, comprising the antibody of
 CC (M1). The present sequence is a murine trichohyalin-related protein,
 CC which was used as the antigen in the method of the invention.
 XX
 XX Sequence 1439 AA;
 SQ
 Query Match 21.1%; Score 329.5; DB 8; Length 1439;
 Best Local Similarity 29.8%; Pred. No. 2.3e-16;
 Matches 108; Conservative 87; Mismatches 103; Indels 65; Gaps 16;
 QY 2 DLEQERRAKK-----LQEQSDLEODRLAK--EKLQEQ-----SDLE 38
 DB 602 ELEQERRRRQQRDRLLEEQFQREHQREARRRDETFOEEBQLQGSRRRQOEREKFTLE 661
 QY 39 QERRAKKLEQEQSDLEQERRAKK--LQEQSDLEQERRAKKLEQEQSDLEODRLAK 95
 DB 662 EERQLTEREEQRRQOEREFQEEBHLQEREKELQECDKRSRQERRQOEREQRLR 721
 QY 96 EKLQEQSDLEQER--RAKELQEQSDLEQER--RAKELQEQSDLEQER--RAK 146
 DB 722 QE-RDQFRREQRRQOERHLEEEQLRDRPFRREQRRLEEEBQLRDRPFRREQRRH 780
 QY 147 EKLQEQSDLEQER--RAKELQEQSDLEQER--RAKELQEQSDLEQERRAKKEL 200
 DB 781 EQLRDRPFRREQRRQOERHLEEEQLRDRPFRREQRRLEEEBQLRDRPFRREQRRH-QERE 839
 QY 201 QEQORLEQKADTKNLERKKEHGDILAEGLYGLLEIPALIPSENERGYIIPHOSSLP 260
 DB 840 EQLRDRPFRR--EQERLLEEEQ--LDRSFRRQ--ELRRDKTHEEERRELE 891
 QY 261 QNRGNSRDSKISITTEKTNR-----SITTVGRRDIKHGLKEKDGSIKP 309
 DB 892 EEQRRGERDLRVEBQLRGQREEQRRQECRKLRELEVQLEERLDRK--LRR 948
 QY 310 EQK 312
 DB 949 EQE 951
 RESULT 10
 ADU77050
 ID ADU77050 standard; protein; 1439 AA.
 XX AC ADU77050;
 XX DT 10-FEB-2005 (first entry)
 XX DE Murine hair growth stimulator related protein.
 XX KW Hair growth stimulation; endocrine-gen.; cardiovascular-gen.;

KW antidiabetic; hepatotropic; antiinflammatory; gastrointestinal-gen.;
 KW muscular-gen.; stimulator of hair growth; cardiovascular disease;
 KW diabetes; liver disease; periodontal disease; gastrointestinal disease;
 KW muscular dystrophy; cholestyrlite.
 OS Mus musculus.
 PN WO2004101610-A1.
 XX 25-NOV-2004.
 XX 14-MAY-2004; 2004WO-JP006909.
 XX 16-MAY-2003; 2003JP-00138567.
 XX 17-NOV-2003; 2003JP-00386185.
 XX (SUME) SUMITOMO ELECTRIC IND LTD.
 XX Hirai Y, Nakajima K;
 XX WPI: 2004-833974/82.
 DR Novel oligopeptide or its modified form useful for enhancing hair growth
 PT or for treating cardiovascular diseases, diabetes, liver diseases,
 PT periodontal disease, intestinal disease or muscular dystrophy.
 XX Disclosure; SEQ ID NO 11; 80pp; Japanese.
 XX
 XX The invention relates to an oligopeptide or its modified form, used for
 CC promoting hair growth, comprising 5-100 amino acid residues. The
 CC invention also relates to a composition comprising the oligopeptide, its
 CC modified form or its salt, and a hair growth stimulant comprising the
 CC oligopeptide, its modified form or its salt. The oligopeptide is useful
 CC for enhancing hair growth or for treating cardiovascular diseases,
 CC diabetes, liver diseases, periodontal disease, intestinal disease,
 CC muscular dystrophy or cholestyrlis. This sequence represents a murine
 CC hair growth stimulation related protein of the invention.
 XX
 XX Sequence 1439 AA;
 SQ
 Query Match 21.1%; Score 329.5; DB 8; Length 1439;
 Best Local Similarity 29.8%; Pred. No. 2.3e-16;
 Matches 108; Conservative 87; Mismatches 103; Indels 65; Gaps 16;
 QY 2 DLEQERRAKK-----LQEQSDLEODRLAK--EKLQEQ-----SDLE 38
 DB 602 ELEQERRRRQQRDRLLEEQFQREHQREARRRDETFOEEBQLQGSRRRQOEREKFTLE 661
 QY 39 QERRAKKLEQEQSDLEQERRAKK--LQEQSDLEQERRAKKLEQEQSDLEODRLAK 95
 DB 662 EERQLTEREEQRRQOEREFQEEBHLQEREKELQECDKRSRQERRQOEREQRLR 721
 QY 96 EKLQEQSDLEQER--RAKELQEQSDLEQER--RAKELQEQSDLEQER--RAK 146
 DB 722 QE-RDQFRREQRRQOERHLEEEQLRDRPFRREQRRLEEEBQLRDRPFRREQRRH 780
 QY 147 EKLQEQSDLEQER--RAKELQEQSDLEQER--RAKELQEQSDLEQERRAKKEL 200
 DB 781 EQLRDRPFRREQRRQOERHLEEEQLRDRPFRREQRRLEEEBQLRDRPFRREQRRH-QERE 839
 QY 201 QEQORLEQKADTKNLERKKEHGDILAEGLYGLLEIPALIPSENERGYIIPHOSSLP 260
 DB 840 EQLRDRPFRR--EQERLLEEEQ--LDRSFRRQ--ELRRDKTHEEERRELE 891
 QY 261 QNRGNSRDSKISITTEKTNR-----SITTVGRRDIKHGLKEKDGSIKP 309
 DB 892 EEQRRGERDLRVEBQLRGQREEQRRQECRKLRELEVQLEERLDRK--LRR 948
 QY 310 EQK 312
 DB 949 EQE 951

DR WPI; 2000-387829/33.
DR N-PSDB; AAA30290.
XX Treating or preventing a disease associated with rhodino virus infection
XX PT in a mammal which includes Kaposi's Sarcoma and Primary Effusion
XX PT Lymphoma.
XX
XX Disclosure; Fig 7; 70pp; English.
XX
XX The present sequence is the Kaposi's sarcoma-associated herpesvirus,
XX (KSHV) latency-associated nuclear antigen (LANA). KSHV is also known as
XX Human Herpes Virus 8 (HHV8) and belongs to the rhadino virus, or gamma-2
XX herpes virus class. The LANA protein is necessary for the efficient
XX persistence of rhadino virus DNA in mammalian cells. Persistent rhadino
XX virus infection is implicated in a variety of diseases e.g. Kaposi's
XX Sarcoma (KS), Primary Effusion Lymphoma (PEL) and multicentric
XX Castleman's disease. In addition, KS is a common malignancy in HIV
XX patients. KSHV persists in host cells in a latent form. One of the few
XX genes expressed from the latent viral DNA is LANA. LANA associates with
XX both human chromosomes and with the rhadino virus cis-acting element
XX (RVCAE), thereby providing a tethering function: the KSHV DNA episome is
XX "tied" to the host chromosomes. This allows the viral DNA to persist in
XX the host cell. The present sequence may be used to screen and identify
XX molecules that inhibit LANA interaction with RVCAE, thereby interfering
XX with the latency cycle of this virus. Potential antiviral treatments for
XX the above mentioned diseases may therefore be based on LANA deregulation.
XX (Updated on 12-SEP-2003 to standardise OS field)
XX
XX Sequence 1162 AA;
SQ
Query Match 21.0%; Score 327.5; DB 3; Length 1162;
Best Local Similarity 25.5%; Pred. No. 2.6e-16;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;
QY 4 EERRAKEKLEQSQSDLEODRLAKELQ--EQQSDLEERRAKEKLEQSQSDLEERRAK 61
DB 654 EQQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQ 713
QY 62 EKLQEQSQSDLEQ--ERRAKEKLEQSQSDLEODRLAKELQEQSQSDLEERRAKEKLEQEQ 119
DB 714 QQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQ 773
QY 120 SDLE--QERRAKE--KLEQSQSDLEQERR--AKEKLEQSQSDLEQERR--AKEKLEQ 167
DB 774 QLELEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLE 833
QY 168 EQQSDL-----EQERRAKEKLEQ--QSDLEQ--ERRAKEKLEQEQRLDQKADTKNLE 219
DB 834 EQQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBE 893
QY 220 RKKEHGDILAEDLYGRLLEIPALELPSENERGYIIPHQSSLPD--NRGNSRDSKEISIT 276
DB 894 EVEEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLE 953
QY 277 EKTNRRESITTNVEG 290
DB 954 VVSTHEQIASSPPG 967
RESULT 13
AAVS8500
ID AAVS8500 standard; protein; 1162 AA.
XX
XX AAVS8500;
XX
XX 06-AUG-2003 (revised)
XX DT 10-APR-2000 (first entry)
XX
XX HHV8 ORF 73 protein, SEQ ID NO:21.
XX
XX HHV8; detection; diagnosis; Kaposi's sarcoma; AIDS; immunogen; antigen.
XX
XX Human herpesvirus 8.

XX
XX Key Location/Qualifiers
XX FH Misc-difference 96
XX PT /label= unknown
XX
XX WO9961909-A2.
XX
XX 02-DEC-1999.
XX
XX 26-MAY-1999; 99WO-US011407.
XX
XX 26-MAY-1998; 98US-0086695P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Pau C;
XX
XX WPI; 2000-097142/08.
XX
XX New methods and compositions for the detection of human herpesvirus.
XX
XX Claim 2; Page 59-62; 68pp; English.
XX
XX Sequences AAVS8480-Y58532 represent immunogenic polypeptides derived from
XX human herpes virus type 8 (HHV8, a gammaherpesvirus). HHV8 plays an
XX important role in the pathogenesis of AIDS-related Kaposi's sarcoma. The
XX invention relates to a novel method of detecting the presence of human
XX herpesvirus 8 in a biological sample using peptides representative of
XX dominant antigenic regions of HHV8. The method comprises contacting one
XX or more isolated, immunogenic HHV8 peptides with an antibody-containing
XX biological sample, and detecting the formation of a complex between the
XX peptide and the antibody. The presence of a peptide-antibody complex
XX indicates the presence of human herpesvirus 8. The detection of HHV8
XX infection can be used to diagnose AIDS-associated Kaposi's sarcoma. The
XX HHV8-specific antibodies are useful therapeutically when for the passive
XX immunisation of a human against HHV8 infection, thereby reducing HHV8
XX related disease. The detection assays are highly specific, sensitive and
XX accurate. Early detection and treatment of Kaposi's sarcoma could
XX diminish the severity of symptoms related to AIDS and the sensitive
XX techniques could reduce erroneous characterisations of skin disorders.
XX Previous assays for HHV8 antibodies such as immunofluorescence assays,
XX immunoblots and enzyme immunoassays lack the sensitivity and accuracy
XX needed for reliable diagnosis of Kaposi's sarcoma. Further advantages of
XX the assays are that reproducible results are obtained and the method is
XX suitable for rapid throughput and screening of samples economically.
XX (Updated on 06-AUG-2003 to correct OS field.)
XX
XX Sequence 1162 AA;
SQ
Query Match 21.0%; Score 327.5; DB 3; Length 1162;
Best Local Similarity 25.5%; Pred. No. 2.6e-16;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;
QY 4 EERRAKEKLEQSQSDLEODRLAKELQ--EQQSDLEERRAKEKLEQSQSDLEERRAK 61
DB 654 EQQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQ 713
QY 62 EKLQEQSQSDLEQ--ERRAKEKLEQSQSDLEODRLAKELQEQSQSDLEERRAKEKLEQEQ 119
DB 714 QQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQSQDEQ 773
QY 120 SDLE--QERRAKE--KLEQSQSDLEQERR--AKEKLEQSQSDLEQERR--AKEKLEQ 167
DB 774 QLELEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLE 833
QY 168 EQQSDL-----EQERRAKEKLEQ--QSDLEQ--ERRAKEKLEQEQRLDQKADTKNLE 219
DB 834 EQQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBE 893
QY 220 RKKEHGDILAEDLYGRLLEIPALELPSENERGYIIPHQSSLPD--NRGNSRDSKEISIT 276
DB 894 EVEEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLEBEQEQLE 953

OY 277 EKTRESITTNVEG 290
DB 954 VVSTHEQIASSPG 967

RESULT 14
AAB62331
ID AAB62331 standard; protein, 1162 AA.
XX
AC AAB62331;
XX
DT 06-AUG-2003 (revised)
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of KSHV tethering protein LANA.
XX
KM Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;
KM Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;
KM KSHV; latency-associated nuclear antigen; LANA.
XX
OS Human herpesvirus 8.
XX
PN WO200125484-A2.
XX
PD 12-APR-2001.
XX
PF 29-SEP-2000; 2000WO-US026908.
XX
PR 01-OCT-1999; 99US-00410399.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Robertson ES, Cotter MA;
XX
DR WPI; 2001-281736/29.
DR N-PSDB; AAF82901.
XX
PT A composition for use in gene therapy comprises an expression vector that
PT includes a nucleic acid sequence encoding a nucleic acid binding protein.
XX
PS Disclosure; Fig 9B; 60pp; English.
XX
CC The invention provides a composition comprising nucleic acid, histone H1
CC protein and expression vector operably encoding a protein suitable
CC for tethering the nucleic acid to the histone H1 protein, where the
CC tethering protein is LANA. The composition is useful in aiding the
CC retention of the viral DNA in the host cell. The viral vector encodes a
CC protein suitable for tethering DNA to histone H1. Methods for screening
CC for compounds which are agonistic or antagonistic for the tethering of
CC viral proteins to histone H1 and DNA binding sites are useful for
CC developing the method of viral transfer. The composition has applications
CC to gene therapy, including the treatment of multiple sclerosis,
CC Parkinson's disease, Huntington disease and diabetes. The present
CC sequence represents the amino acid sequence of the Kaposi's sarcoma
CC associated herpesvirus (human herpesvirus 8) latency-associated nuclear
CC antigen (LANA), which acts as a tethering protein. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 1162 AA;

Query Match 21.0%; Score 327.5; DB 4; Length 1162;
Best Local Similarity 25.5%; Pred. No. 2.6e-16;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;

OY 4 EQRRAKELQEQSDLEQDRLAKEKIQ--EQQSDLEQRRAKELQEQSDLEQRRAK 61
DB 654 EQQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQ 713
OY 62 EKLQEQSQDLEQ--ERRAKELQEQSDLEQDRLAKEKIQEQSDLEQRRAKELQEQ 119
DB 714 QDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQ 773
OY 120 SDLE-QERRAKE---KLEQEQSDLEQERR---AKEKIQEQSDLEQERR---AKEKIQ 167

DB 774 QLELDQELQELQEQELQELQELQELQELQELQELQELQELQELQELQELQEL 833
OY 168 EQQSDLE-----EQRRRAKELQEQ--QSDLEQ--ERRAKELQEQQDRLEQRKADTKKNLE 219
DB 834 EQQELQEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVE 893
OY 220 RKKEHGDILAEADLYGRLEIPALIEPSENENGYIYIPHOSSLPQD---NRGNSRDSKEISII 276
DB 894 EVEEQEQLQELQEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQ 953
OY 277 EKTRESITTNVEG 290
DB 954 VVSTHEQIASSPG 967

RESULT 15
ABB05621
ID ABB05621 standard; protein, 1162 AA.
XX
AC ABB05621;
XX
DT 25-APR-2002 (first entry)
XX
DE Kaposi's sarcoma-associated herpesvirus LANA protein.
XX
KM Kaposi's sarcoma-associated herpesvirus; KSHV; LANA; RVCAE; PEL;
KM KSHV terminal repeat; rhadino virus cis acting element; episome;
KM primary effusion lymphoma; latency-associated nuclear antigen;
KM gene therapy; gene transfer.
XX
OS Human herpesvirus 8.
XX
PN US6322792-B1.
XX
PD 27-NOV-2001.
XX
PF 21-APR-1999; 99US-00298568.
XX
PR 19-NOV-1998; 98US-0109422P.
XX
PA (KIEF/) KIEF E D.
XX
PI Kieff ED, Ballesgas ME, Kaye KM;
XX
DR WPI; 2002-153769/20.
DR N-PSDB; ABA93487.
XX
PT System for episomal retention of plasmids in mammalian cells, useful in
PT gene therapy, comprises rhadinoviral LANA and RVCAE sequences.
XX
PS Disclosure; Fig 7; 27pp; English.
XX
CC The present invention describes a system (A) for maintaining a plasmid as
CC an episome in mammalian cells, comprising the rhadinoviral sequence LANA
CC (latency-associated nuclear antigen) of 3469 base pairs (see ABA93487,
CC S1) expressed in the cell, and the rhadinoviral sequence RVCAE
CC (rhadinoviral cis-acting element) of 801 base pairs (see ABA93488, S2)
CC present in the plasmid. Also described is a method for maintaining a
CC closed circular DNA in a cell by expressing (S1) in the cells and having
CC (S2) as a cis-acting and maintenance sequence in the DNA. (A) is
CC particularly used in gene therapy (or other gene transfer applications)
CC that uses mammalian cells in which LANA is expressed. (A) improves
CC persistence of gene therapy vectors in cells. The present sequence
CC represents Kaposi's sarcoma-associated herpesvirus (KSHV, also called
CC human herpesvirus 8) LANA protein, which is used in the exemplification
CC of the present invention
XX
SQ Sequence 1162 AA;

Query Match 21.0%; Score 327.5; DB 5; Length 1162;
Best Local Similarity 25.5%; Pred. No. 2.6e-16;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:16:04 ; Search time 31.3811 Seconds

(without alignments)
832.525 Million cell updates/sec

Title: US-09-837-344-31

Sequence: 1 SDLEGERRAKEKKEEQSDL.....GHLEKKDSIRKEQKEDKS 316

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/prodata/1/1aa/5 COMB.pep:*

2: /cgn2_6/prodata/1/1aa/6 COMB.pep:*

3: /cgn2_6/prodata/1/1aa/H COMB.pep:*

4: /cgn2_6/prodata/1/1aa/PCUS COMB.pep:*

5: /cgn2_6/prodata/1/1aa/RE COMB.pep:*

6: /cgn2_6/prodata/1/1aa/backfil1a1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	316	2	US-08-098-327E-31 Sequence 31, Appl
2	1563	100.0	316	2	US-08-462-625-31 Sequence 31, Appl
3	603	38.6	117	2	US-08-098-327E-20 Sequence 20, Appl
4	603	38.6	117	2	US-08-462-625-20 Sequence 20, Appl
5	554	35.4	107	2	US-08-098-327E-19 Sequence 19, Appl
6	554	35.4	107	2	US-08-462-625-19 Sequence 19, Appl
7	344	22.0	608	2	US-09-270-767-32937 Sequence 32937, A
8	344	22.0	608	2	US-09-270-767-48154 Sequence 48154, A
9	327.5	21.0	1162	1	US-08-728-323A-2 Sequence 2, Appl
10	327.5	21.0	1162	2	US-09-298-568-2 Sequence 2, Appl
11	327.5	21.0	1162	2	US-09-410-359-2 Sequence 2, Appl
12	327.5	21.0	1162	2	US-09-894-273-2 Sequence 2, Appl
13	326	20.9	611	2	US-09-216-393B-81 Sequence 81, Appl
14	317.5	20.3	1898	1	US-08-056-200-94 Sequence 94, Appl
15	317.5	20.3	1898	1	US-08-800-644-94 Sequence 94, Appl
16	317.5	20.3	1898	1	US-09-538-092-1280 Sequence 1280, Ap
17	301	19.3	1180	2	US-09-543-681A-6436 Sequence 6436, Ap
18	288	18.4	650	2	US-10-104-047-3636 Sequence 3636, Ap
19	278	17.8	280	2	US-10-104-047-3063 Sequence 3063, Ap
20	277	17.7	735	2	US-10-164-595-80 Sequence 80, Appl
21	277	17.7	784	2	US-10-164-595-79 Sequence 79, Appl
22	277	17.7	843	2	US-10-164-595-54 Sequence 54, Appl
23	275	17.6	779	2	US-10-164-595-56 Sequence 56, Appl
24	272	17.4	1564	2	US-10-144-198-2 Sequence 2, Appl
25	272	17.4	1564	2	US-10-144-198-4 Sequence 4, Appl
26	270.5	17.3	919	2	US-09-949-016-6954 Sequence 6954, Ap
27	268	17.1	798	2	US-10-104-047-3335 Sequence 3335, Ap

28	265	17.0	676	2	US-09-270-767-44900	Sequence 44900, A
29	260	16.6	1402	2	US-09-248-796A-14503	Sequence 14503, A
30	259	16.6	613	2	US-09-543-681A-6012	Sequence 6012, Ap
31	255.5	16.3	1960	2	US-09-538-092-10772	Sequence 1077, Ap
32	255.5	16.3	1960	2	US-09-538-092-10772	Sequence 1077, Ap
33	255	16.3	683	6	5210183-3	Patent No. 5210183
34	253	16.2	2101	1	US-08-466-390-4	Sequence 4, Appl
35	253	16.2	2101	1	US-08-470-950-4	Sequence 4, Appl
36	253	16.2	2101	1	US-08-467-781-4	Sequence 4, Appl
37	253	16.2	2101	1	US-08-195-487-4	Sequence 4, Appl
38	253	16.2	2101	1	US-08-483-924-4	Sequence 4, Appl
39	253	16.2	2101	2	US-09-452-294-1	Sequence 32, Appl
40	253	16.2	2101	2	US-09-296-662-32	Sequence 32, Appl
41	253	16.2	1976	4	PCT-US93-06160-4	Sequence 4, Appl
42	250.5	16.0	1976	2	US-09-538-092-1078	Sequence 1078, Ap
43	250	16.0	2107	2	US-09-949-016-7646	Sequence 7646, Ap
44	250	16.0	2107	2	US-09-949-016-7647	Sequence 7647, Ap
45	250	16.0	2115	2	US-09-296-662-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-08-098-327E-31
; Sequence 31, Application US/08098327E
; Patent No. 6270771
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRULHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098,327E
; FILING DATE: 24-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; US-08-098-327E-31

Query Match 100.0%; Score 1563; DB 2; Length 316;

Best Local Similarity 100.0%; Pred. No. 1.6e-119; Indels 0; Gaps 0;
Matches 316; Conservative 0; Mismatches 0;

QY 1 SDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRA 60
Db 1 SDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRA 60
QY 61 KEKLEQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSD 120
Db 61 KEKLEQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSD 120
QY 121 DLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 180
Db 121 DLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 180
QY 181 EKLQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 240
Db 181 EKLQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 240
QY 241 IELPENERGYIIPHQSLPDNRGNSRDSKEISIEKTNRSTTTNVEGRDIIKHGLE 300
Db 241 IELPENERGYIIPHQSLPDNRGNSRDSKEISIEKTNRSTTTNVEGRDIIKHGLE 300
QY 301 EKKDGSIRPEQKEDKS 316
Db 301 EKKDGSIRPEQKEDKS 316

RESULT 2

US-08-462-625-31
Sequence 31, Application US/08462625
Patent No. 6319502

GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: DRUGS, Pierre
TITLE OF INVENTION: HEPATIC STAGES OF P. PALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-31

Query Match 100.0%; Score 1563; DB 2; Length 316;
Best Local Similarity 100.0%; Pred. No. 1.6e-119; Indels 0; Gaps 0;
Matches 316; Conservative 0; Mismatches 0;

QY 1 SDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRA 60
Db 1 SDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRA 60
QY 61 KEKLEQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSD 120
Db 61 KEKLEQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSD 120
QY 121 DLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 180
Db 121 DLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 180
QY 181 EKLQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 240
Db 181 EKLQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 240
QY 241 IELPENERGYIIPHQSLPDNRGNSRDSKEISIEKTNRSTTTNVEGRDIIKHGLE 300
Db 241 IELPENERGYIIPHQSLPDNRGNSRDSKEISIEKTNRSTTTNVEGRDIIKHGLE 300
QY 301 EKKDGSIRPEQKEDKS 316
Db 301 EKKDGSIRPEQKEDKS 316

RESULT 3

US-08-098-327E-20
Sequence 20, Application US/08098327E
Patent No. 6270771

GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: DRUGS, Pierre
TITLE OF INVENTION: HEPATIC STAGES OF P. PALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-20

Query Match 38.6%; Score 603; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.9e-42;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 200 LQEQQRDLQQRKADTKNLERKKEHGDILADLYGRLEIPALIELPSENERGYIIPHOSSL 259
Db 1 LQEQQRDLQQRKADTKNLERKKEHGDILADLYGRLEIPALIELPSENERGYIIPHOSSL 60

Qy 260 PQRNGSRDSKESIIETKNTRESITTNVEGRDIIHGHEKKDGSIKPEQKEDKS 316
Db 61 PQRNGSRDSKESIIETKNTRESITTNVEGRDIIHGHEKKDGSIKPEQKEDKS 117

RESULT 4
US-08-462-625-20
Sequence 20, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-20

Query Match 38.6%; Score 603; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.9e-42;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 200 LQEQQRDLQQRKADTKNLERKKEHGDILADLYGRLEIPALIELPSENERGYIIPHOSSL 259
Db 1 LQEQQRDLQQRKADTKNLERKKEHGDILADLYGRLEIPALIELPSENERGYIIPHOSSL 60

Qy 260 PQRNGSRDSKESIIETKNTRESITTNVEGRDIIHGHEKKDGSIKPEQKEDKS 316
Db 61 PQRNGSRDSKESIIETKNTRESITTNVEGRDIIHGHEKKDGSIKPEQKEDKS 117

RESULT 5
US-08-098-327E-19
Sequence 19, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-19

Query Match 35.4%; Score 554; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 4.2e-38;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 210 RKADTKNLERKKEHGDILADLYGRLEIPALIELPSENERGYIIPHOSSLPQRNGSRD 269
Db 1 RKADTKNLERKKEHGDILADLYGRLEIPALIELPSENERGYIIPHOSSLPQRNGSRD 60

QY 270 SKEISIEKTNRESITTNVEGRDIIHKHLEEKDGSIKPEOKEDKS 316
Db 61 SKEISIEKTNRESITTNVEGRDIIHKHLEEKDGSIKPEOKEDKS 107

RESULT 6

US-08-462-625-19
Sequence 19, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Matchie
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-19

Query Match 35.4%; Score 554; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 4, 2e-38;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 210 RKADTKKNIERKKEHGDIIAEDLYGRLEIPALPSENERGYIIPHOSSIPQDNRRNSRD 269
Db 1 RKADTKKNIERKKEHGDIIAEDLYGRLEIPALPSENERGYIIPHOSSIPQDNRRNSRD 60
QY 270 SKEISIEKTNRESITTNVEGRDIIHKHLEEKDGSIKPEOKEDKS 316
Db 61 SKEISIEKTNRESITTNVEGRDIIHKHLEEKDGSIKPEOKEDKS 107

RESULT 7
US-09-270-767-32937

Sequence 32937, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 32937
LENGTH: 608
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-32937

Query Match 22.0%; Score 344; DB 2; Length 608;
Best Local Similarity 31.7%; Pred. No. 3, 5e-20;
Matches 71; Conservative 77; Mismatches 68; Indels 8; Gaps 4;

QY 4 EERRAKEXLQEOQSDLEQDRLAKEXLQEOQSDLEQERRAKEXLQEOQSDLEQERRAKEX 63
Db 381 EERRKKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEER 440
QY 64 LQEOQSDLEQERRAKEXLQEOQSDLEQDRLAKEXLQEOQSDLEQERRAKEXLQEOQSDLE 123
Db 441 REBAERK-EERRKAERKKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKE 496
QY 124 QERRA-KEKLOEQSDLEQERRAKEXLQEOQSDLEQERRAKEXLQEOQSDLEQERRAKEX 182
Db 497 EERRKDEERRERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEER 556
QY 183 LQEOQSDLEQERRAKEXLQEOQSDLEQ--RKADTKKNIERKKE 223
Db 557 RDEERRERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKE 600

RESULT 8
US-09-270-767-48154
Sequence 48154, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 48154
LENGTH: 608
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-48154

Query Match 22.0%; Score 344; DB 2; Length 608;
Best Local Similarity 31.7%; Pred. No. 3, 5e-20;
Matches 71; Conservative 77; Mismatches 68; Indels 8; Gaps 4;
QY 4 EERRAKEXLQEOQSDLEQDRLAKEXLQEOQSDLEQERRAKEXLQEOQSDLEQERRAKEX 63
Db 381 EERRKKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEER 440
QY 64 LQEOQSDLEQERRAKEXLQEOQSDLEQDRLAKEXLQEOQSDLEQERRAKEXLQEOQSDLE 123
Db 441 REBAERK-EERRKAERKKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKE 496
QY 124 QERRA-KEKLOEQSDLEQERRAKEXLQEOQSDLEQERRAKEXLQEOQSDLEQERRAKEX 182
Db 497 EERRKDEERRERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEER 556
QY 183 LQEOQSDLEQERRAKEXLQEOQSDLEQ--RKADTKKNIERKKE 223

Db 557 RDEERRRERREERREERREERREERREERREERREERREERREERRE 600

RESULT 9

US-08-728-323A-2
; Sequence 2, Application US/06728323A
; Patent No. 5946676
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Rueso, James J.
; APPLICANT: Edelman, Jaidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,323A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/52268/JPM/MSC/SKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-728-323A-2

Query Match 21.0%; Score 327.5; DB 1; Length 1162;
Best Local Similarity 25.5%; Pred. No. 1.6e-18;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;
QY 4 EERRAKKLEQEQSDLEQDRRLAKKLEQ--EQSDLEQERRAKKLEQEQSDLEQERRAK 61
Db EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQ 713
QY 62 EKLQEQSDLEQ--ERRAKKLEQEQSDLEQDRRLAKKLEQEQSDLEQERRAKKLEQEQ 119
Db EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQ 713
QY 714 QQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQ 773
Db 120 SDLE-QERRAKE--KLEQEQSDLEQERR--AKKLEQEQSDLEQERR--AKKLEQ 167
QY 774 QELDQEQLEEQEQLEEQEQLEEQEQLEEQEQLEEQEQLEEQEQLEEQEQLEEQEQLE 833
Db 168 EQQSDLE-----EQERRAKKLEQEQ--QSDLEQ--ERRAKKLEQEQORDLEQKADTKKYLE 219
QY 834 EQEQLEEQEQVEEQEQVEEQEQVEEQEQVEEQEQVEEQEQVEEQEQVEEQEQVEEQEQLE 893
Db 220 RKKEHGDILADLYGRLEIPALIEPSENERGYIIPHOSSLPOD--NRGNSRDSKEISII 276
QY 894 EEEQEQEQLEVEEQEQEQLEVEEQEQEQVEEQEQEQVEEQEQEQVEEQEQEQVEEQEQ 953
QY 277 EKTNRSEITTNVEG 290

Db 954 VVSTHEQIASSPG 967

RESULT 10

US-09-298-568-2
; Sequence 2, Application US/09298568
; Patent No. 6322792
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Balleskas, Mary E.
; APPLICANT: Kaye, Kenneth M.
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/298,568
; CURRENT FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,422
; EARLIER FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
; US-09-298-568-2

Query Match 21.0%; Score 327.5; DB 2; Length 1162;
Best Local Similarity 25.5%; Pred. No. 1.6e-18;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;

QY 4 EERRAKKLEQEQSDLEQDRRLAKKLEQ--EQSDLEQERRAKKLEQEQSDLEQERRAK 61
Db EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQ 713
QY 62 EKLQEQSDLEQ--ERRAKKLEQEQSDLEQDRRLAKKLEQEQSDLEQERRAKKLEQEQ 119
Db EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQ 773
QY 714 QQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQ 773
Db 120 SDLE-QERRAKE--KLEQEQSDLEQERR--AKKLEQEQSDLEQERR--AKKLEQ 167
QY 774 QELDQEQLEEQEQLEEQEQLEEQEQLEEQEQLEEQEQLEEQEQLEEQEQLEEQEQLE 833
Db 168 EQQSDLE-----EQERRAKKLEQEQ--QSDLEQ--ERRAKKLEQEQORDLEQKADTKKYLE 219
QY 834 EQEQLEEQEQVEEQEQVEEQEQVEEQEQVEEQEQVEEQEQVEEQEQVEEQEQVEEQEQLE 893
Db 220 RKKEHGDILADLYGRLEIPALIEPSENERGYIIPHOSSLPOD--NRGNSRDSKEISII 276
QY 894 EEEQEQEQLEVEEQEQEQLEVEEQEQEQVEEQEQEQVEEQEQEQVEEQEQEQVEEQEQ 953
QY 277 EKTNRSEITTNVEG 290
Db 954 VVSTHEQIASSPG 967
RESULT 11
US-09-410-399-2
; Sequence 2, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
; APPLICANT: Robertson, Erle S.
; APPLICANT: Cotter, Murray A.
; TITLE OF INVENTION: Method to Inhibit or Enhance the Binding of Viral DNA
; FILE REFERENCE: UM-03778
; CURRENT APPLICATION NUMBER: US/09/410,399
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1162
; TYPE: PRT

```

; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-410-389-2
Query Match      21.0%; Score 327.5; DB 2; Length 1162;
Best Local Similarity 25.5%; Pred. No. 1.6e-18;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;

QY 4 EERRAKKLEQEQSDLEQDRLAKELQ--EQQSDLEERRAKKLEQEQSDLEERRAK
DB 654 EQQDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQ
QY 62 EKLQEQSDLEQ--ERRAKKLEQEQSDLEQDRLAKELQEQSDLEQERRAKKLEQEQ
DB 714 QQDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQ
QY 120 SDLE-QERRAKE--KLQEQSDLEQERR--AKKLEQEQSDLEQERR--AKKLEQ
DB 774 QELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELE
QY 168 EQQSDLE-----EQERRAKKLEQEQ-QSDLEQ-ERRAKKLEQEQORDLEQKADTKKLE
DB 834 EQEQELEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQELE
QY 220 RKKEGDIADLYGRLEIPALPELPSENERGYIIPHOSSLPOD--NRGNSRDSKEISII
DB 894 EYEEQEQELEVEEQEQELEVEEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ
QY 277 EKTNRRESITTTNVEG 290
DB 954 VVSTHEQIASSPPG 967

RESULT 12
US-09-894-273-2
; Sequence 2, Application US/09894273
; Patent No. 6756203
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Ballester, Mary E.
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/894,273
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/109,422
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1162
; TYPE: PR
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-894-273-2
Query Match      21.0%; Score 327.5; DB 2; Length 1162;
Best Local Similarity 25.5%; Pred. No. 1.6e-18;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;

QY 4 EERRAKKLEQEQSDLEQDRLAKELQ--EQQSDLEERRAKKLEQEQSDLEQERRAK
DB 654 EQQDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQ
QY 62 EKLQEQSDLEQ--ERRAKKLEQEQSDLEQDRLAKELQEQSDLEQERRAKKLEQEQ
DB 714 QQDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQDDEQEQ
QY 120 SDLE-QERRAKE--KLQEQSDLEQERR--AKKLEQEQSDLEQERR--AKKLEQ
DB 774 QELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELE
QY 168 EQQSDLE-----EQERRAKKLEQEQ-QSDLEQ-ERRAKKLEQEQORDLEQKADTKKLE
DB 834 EQQSDLE-----EQERRAKKLEQEQ-QSDLEQ-ERRAKKLEQEQORDLEQKADTKKLE
QY 219
```

```

DB 834 EQEQELEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQELE
QY 220 RKKEGDIADLYGRLEIPALPELPSENERGYIIPHOSSLPOD--NRGNSRDSKEISII
DB 894 EYEEQEQELEVEEQEQEQELEVEEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ
QY 277 EKTNRRESITTTNVEG 290
DB 954 VVSTHEQIASSPPG 967

RESULT 13
US-09-216-393B-81
; Sequence 81, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
; APPLICANT: Milnausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDI PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 81
; LENGTH: 611
; TYPE: PR
; ORGANISM: Toxoplasma gondii
US-09-216-393B-81
Query Match      20.9%; Score 326; DB 2; Length 611;
Best Local Similarity 26.1%; Pred. No. 1e-18;
Matches 89; Conservative 99; Mismatches 99; Indels 54; Gaps 12;

QY 4 EQERR--AKKLEQEQSDLEQDRLAKELQEQSDLEQERR--AKKLEQEQSDLEQEQ
DB 292 EERRRAREEKEKERRRQEQEERERRRVEEKEKERRRQEQEERERRRVEEKEKARRQEQEERERR
QY 58 RRAKELQEQSDLEQERR--AKKLEQEQSDLEQDRLAKELQEQSDLEQERRAKKLE
DB 352 RRVEEKARRQEQEERERRRVEEKEKARRQEQEERERRRVEEKEKARRQEQEERERRRVE
QY 114 KLEQEQSDLEQERRAK--EKLQEQSDLEQEQ-RRAKKLEQEQSDLEQERRAKKLEQEQ
DB 412 EKKARRQEQEERERRRVEEKEKARRQEQEERERRRVEEKEKARRQEQEERERRRVEEKE
QY 169 ---QQSDLEQERRAKKLEQEQSDLEQERR--AKKLEQEQORDLEQKADTKKLE
DB 472 RERQEQEERERRRVEEKEKERRRQEQEERERRRVEEKEKERRRQEQEERERRRVEEKEKARRQEQ
QY 222 KEHGDILADLYGRLEIPALPELPSENERGYIIPHOSSLPODNRGNSRDSK-EISIIKTN
DB 532 EEEG-----RERQEQ--EERQEQEERERRRQEQEERERRRQEQEERERRRQEQEERERRR
QY 281 RESITTT-NVEGRDII-----KGHLEKKKGSIKPEOK 312
DB 570 AESTVTKKEQGERRHLLTOPMERKTLHAFEESSRTVLQER 610

RESULT 14
US-08-056-200-94
; Sequence 94, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichogyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Knobbe, Martens, Olson & Bear
;; STREET: 620 Newport Center Drive, Sixteenth Floor
;; CITY: Newport Beach
;; STATE: CA
;; COUNTRY: U.S.A.
;; ZIP: 92660
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/056,200
;; FILING DATE: 30-APR-1993
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fedrick, Michael F.
;; REGISTRATION NUMBER: 36,799
;; REFERENCE/DOCKET NUMBER: NIH054,001A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (714) 760-0404
;; TELEFAX: (714) 760-9502
;; INFORMATION FOR SEQ ID NO: 94:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1898 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-056-200-94

Query Match 20.3%; Score 317.5; DB 1; Length 1898;
Best Local Similarity 29.7%; Pred. No. 1.9e-17;
Matches 105; Conservative 77; Mismatches 116; Indels 55; Gaps 14;

```
Oy 4 EERRAKKLEQEQSDLEODRLAKKLEQEQSDLEQERRAKKLEQEQSDL-----E 55
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 458 EERR-EGRLKREGE--RDWLKREETERH---EQERRKQQLKRDQEEERRERWLKLE 512
Oy 56 QERRAKKLEQEQSDLEQERRAKKLEQEQSDLEODRLAKKLEQEQSDLEQERRAKKLE 115
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 513 EERRRQERRRQQLRREQERRRQQLKRE--EERLQQLRSLRQQLRREQERRLEQL 569
Oy 116 -QEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSD-L 173
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 570 KREERKRLQERRRQQLRREQERRRQQLKREERKQQLKREQERRLEQRLKREEVERL 629
Oy 174 EERRAKKLEQEQ-----SDLEQERRAKKLEQEQORDLEQ--KADTKKULER 220
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 630 EQERRRDLRLKREPEERRRHLLKSEQERRRHQQLRREQERRRQQLKREEEERLEQ 689
Oy 221 --KKEGDLAEDLYGLE-----IPATLPSENERGYIIPQSSL---PODN 263
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 690 RLKREHEERRRQQLRREQERRRQQLRREQERRRQQLRREQERRRQQLRREQERRRQQL 746
Oy 264 RGNRDSKEISIIKTNRESITTVNVEGRDILHKLLEKKDGSIKPEKEDKS 316
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 747 RAAPQOE-----EKRRRESELQWQERRRAHQEQEERQDRDFTWQQAEEKS 795
```

RESULT 15
US-08-800-644-94
; Sequence 94, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methode of Using Same
; NUMBER OF SEQUENCES: 117

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Knobbe, Martens, Olson & Bear
;; STREET: 620 Newport Center Drive, Sixteenth Floor
;; CITY: Newport Beach
;; STATE: CA
;; COUNTRY: U.S.A.
;; ZIP: 92660
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/800,644
;; FILING DATE: 14-FEB-1997
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/056,200
;; FILING DATE: 30-APR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fedrick, Michael F.
;; REGISTRATION NUMBER: 36,799
;; REFERENCE/DOCKET NUMBER: NIH054,001A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (714) 760-0404
;; TELEFAX: (714) 760-9502
;; INFORMATION FOR SEQ ID NO: 94:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1898 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-800-644-94

Query Match 20.3%; Score 317.5; DB 1; Length 1898;
Best Local Similarity 29.7%; Pred. No. 1.9e-17;
Matches 105; Conservative 77; Mismatches 116; Indels 55; Gaps 14;

```
Oy 4 EERRAKKLEQEQSDLEODRLAKKLEQEQSDLEQERRAKKLEQEQSDL-----E 55
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 458 EERR-EGRLKREGE--RDWLKREETERH---EQERRKQQLKRDQEEERRERWLKLE 512
Oy 56 QERRAKKLEQEQSDLEQERRAKKLEQEQSDLEODRLAKKLEQEQSDLEQERRAKKLE 115
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 513 EERRRQERRRQQLRREQERRRQQLKRE--EERLQQLRSLRQQLRREQERRLEQL 569
Oy 116 -QEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSD-L 173
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 570 KREERKRLQERRRQQLRREQERRRQQLKREERKQQLKREQERRLEQRLKREEVERL 629
Oy 174 EERRAKKLEQEQ-----SDLEQERRAKKLEQEQORDLEQ--KADTKKULER 220
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 630 EQERRRDLRLKREPEERRRHLLKSEQERRRHQQLRREQERRRQQLKREEEERLEQ 689
Oy 221 --KKEGDLAEDLYGLE-----IPATLPSENERGYIIPQSSL---PODN 263
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 690 RLKREHEERRRQQLRREQERRRQQLRREQERRRQQLRREQERRRQQLRREQERRRQQL 746
Oy 264 RGNRDSKEISIIKTNRESITTVNVEGRDILHKLLEKKDGSIKPEKEDKS 316
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 747 RAAPQOE-----EKRRRESELQWQERRRAHQEQEERQDRDFTWQQAEEKS 795
```

Search completed: December 7, 2005, 07:22:12
Job time : 32.3811 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:18:34 ; Search time 82.4665 Seconds
(without alignments)
1601.063 Million cell updates/sec

Title: US-09-837-344-31

Sequence: 1 SDLEQERRAKEXLQEQQSD.....GHLEKKDSIKPEQKEDKS 316

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main:*

- 1: /cgn2_6/ptcdat/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptcdat/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptcdat/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptcdat/1/pubpaa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/ptcdat/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptcdat/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	316	3	US-09-837-344-31 Sequence 31, Appl
2	1563	100.0	316	3	US-09-900-963-31 Sequence 31, Appl
3	1494	95.6	1909	5	US-10-732-923-3341 Sequence 3341, Ap
4	603	38.6	117	3	US-09-837-344-20 Sequence 20, Appl
5	603	38.6	117	3	US-09-900-963-20 Sequence 20, Appl
6	554	35.4	107	3	US-09-837-344-19 Sequence 19, Appl
7	554	35.4	107	3	US-09-900-963-19 Sequence 19, Appl
8	355	22.7	1003	5	US-10-732-923-3319 Sequence 3319, Ap
9	329.5	21.1	1439	5	US-10-754-079-157 Sequence 157, App
10	328	21.0	1407	5	US-10-732-923-3359 Sequence 3359, Ap
11	327.5	21.0	1162	3	US-09-894-273-2 Sequence 2, Appl1
12	327.5	21.0	1162	3	US-10-294-804-2 Sequence 2, Appl1
13	327.5	21.0	1162	4	US-10-194-046-2 Sequence 2, Appl1
14	326	20.9	611	3	US-09-326-193-81 Sequence 81, Appl
15	326	20.9	611	4	US-10-321-856-81 Sequence 81, Appl
16	321	20.5	554	6	US-11-097-143-5154 Sequence 5154, Ap
17	320	20.5	350	4	US-10-101-487-58 Sequence 58, Appl
18	320	20.5	350	5	US-10-939-988-58 Sequence 58, Appl
19	311	19.9	621	4	US-10-108-260A-4409 Sequence 4409, Ap
20	310.5	19.9	1549	5	US-10-732-923-3358 Sequence 3358, Ap
21	307	19.6	240	4	US-10-101-487-75 Sequence 75, Appl
22	307	19.6	240	5	US-10-939-988-75 Sequence 75, Appl
23	306	19.6	2274	6	US-11-097-143-2763 Sequence 2763, Ap
24	289	18.5	923	5	US-10-450-763-38959 Sequence 38959, A
25	288	18.4	650	4	US-10-104-047-3636 Sequence 3636, Ap
26	288	18.4	949	5	US-10-450-763-45867 Sequence 45867, A
27	284	18.2	1715	6	US-11-097-143-1059 Sequence 1059, Ap

28	282	18.0	466	4	US-10-108-260A-4344 Sequence 4344, Ap
29	281	18.0	658	6	US-11-097-143-23688 Sequence 23688, A
30	281	18.0	800	6	US-11-097-143-41169 Sequence 41169, A
31	280.5	17.9	345	4	US-10-029-386-34255 Sequence 34255, A
32	278	17.8	280	4	US-10-104-047-3063 Sequence 3063, Ap
33	277	17.7	735	4	US-10-144-194A-109 Sequence 109, App
34	277	17.7	735	5	US-10-491-566-109 Sequence 109, App
35	277	17.7	735	5	US-10-717-665-80 Sequence 80, Appl
36	277	17.7	784	4	US-10-144-194A-108 Sequence 108, App
37	277	17.7	784	5	US-10-491-566-108 Sequence 108, App
38	277	17.7	784	5	US-10-717-665-79 Sequence 79, Appl
39	277	17.7	843	4	US-10-144-194A-84 Sequence 84, Appl
40	277	17.7	843	5	US-10-491-566-84 Sequence 84, Appl
41	277	17.7	843	5	US-10-717-665-54 Sequence 54, Appl
42	275	17.6	779	4	US-10-144-194A-86 Sequence 86, Appl
43	275	17.6	779	5	US-10-491-566-86 Sequence 86, Appl
44	275	17.6	779	5	US-10-717-665-56 Sequence 56, Appl
45	273	17.5	507	4	US-10-425-114-47285 Sequence 47285, A

ALIGNMENTS

RESULT 1
US-09-837-344-31
Sequence 31, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 31:

Db 1601 EKQEQORDLEGRKAKKEQEQORDLEGRKADTKNLERKKEHGVLAEDLYGRLEIPA 1660
QY 241 IELPSENERGYIIPHQSSLPODNRGNSRDSKSEISIEKTNRESITTVNVEGRDIIHKHGLE 300
Db 1661 IELPSENERGYIIPHQSSLPODNRGNSRDSKSEISIEKTNRESITTVNVEGRDIIHKHGLE 1720
QY 301 EKKDGSIKPEQKEDKS 316
Db 1721 EKKDGSIKPEQKEDKS 1736

RESULT 4
US-09-837-344-20
; Sequence 20, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-837-344-20

Query Match 38.6%; Score 603; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 1,3e-28;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 LOEQORDLEGRKADTKNLERKKEHGVLAEDLYGRLEIPAIELPSENERGYIIPHQSSL 259
Db 1 LOEQORDLEGRKADTKNLERKKEHGVLAEDLYGRLEIPAIELPSENERGYIIPHQSSL 60
QY 260 PODNRGNSRDSKSEISIEKTNRESITTVNVEGRDIIHKHGLEKKGDSIKPEQKEDKS 316

Db 61 PODNRGNSRDSKSEISIEKTNRESITTVNVEGRDIIHKHGLEKKGDSIKPEQKEDKS 117

RESULT 5
US-09-900-963-20
; Sequence 20, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,963
; FILING DATE: 10-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/098,327
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-900-963-20

Query Match 38.6%; Score 603; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 1,3e-28;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 LOEQORDLEGRKADTKNLERKKEHGVLAEDLYGRLEIPAIELPSENERGYIIPHQSSL 259
Db 1 LOEQORDLEGRKADTKNLERKKEHGVLAEDLYGRLEIPAIELPSENERGYIIPHQSSL 60
QY 260 PODNRGNSRDSKSEISIEKTNRESITTVNVEGRDIIHKHGLEKKGDSIKPEQKEDKS 316
Db 61 PODNRGNSRDSKSEISIEKTNRESITTVNVEGRDIIHKHGLEKKGDSIKPEQKEDKS 117

RESULT 6
US-09-837-344-19
; Sequence 19, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine

DRUIHE, Pierre
TITLE OF INVENTION: HEPATIC SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-837-344-19
Query Match 35.4%; Score 554; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 9.2e-26;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 210 RKADTKKNIKKERKKGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 269
DB 1 RKADTKKNIKKERKKGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 60
QY 270 SKEISIIKTNRESITTNVEGRDIIKGLHEKKDGSIKPEOKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIKGLHEKKDGSIKPEOKEDKS 107
RESULT 7
US-09-900-963-19
Sequence 19, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUIHE, Pierre
TITLE OF INVENTION: HEPATIC SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404

CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-900-963-19
Query Match 35.4%; Score 554; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 9.2e-26;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 210 RKADTKKNIKKERKKGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 269
DB 1 RKADTKKNIKKERKKGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 60
QY 270 SKEISIIKTNRESITTNVEGRDIIKGLHEKKDGSIKPEOKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIKGLHEKKDGSIKPEOKEDKS 107
RESULT 8
US-10-732-923-3319
Sequence 3319, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52796)C
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 10/310,154
PRIOR FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 3319
LENGTH: 1003
TYPE: PRT
ORGANISM: Kapoisi's sarcoma-associated herpesvirus
US-10-732-923-3319
Query Match 22.7%; Score 355; DB 5; Length 1003;
Best Local Similarity 25.3%; Pred. No. 3.8e-13;
Matches 79; Conservative 116; Mismatches 63; Indels 54; Gaps 14;
QY 4 EQERRAKELQEQGD--LEQDRLAKELQEQSDLEQERRAKELQEQSDLEQERRAK 61

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:21:30 ; Search time 7.29792 Seconds

(without alignments)
241.806 Million cell updates/sec

Title: US-09-837-344-31

Perfect score: 1563
Sequence: 1 SDLEQERRAKEKLOEQOSDL.....GHLEKKGDSIKPEQKDKS 316

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 32527 seqs, 5584426 residues

Total number of hits satisfying chosen parameters: 32527

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA New:*

- 1: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US66_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	241	15.4	793	7	US-11-060-914-2 Sequence 2, Appli
2	235	15.0	885	6	US-10-793-626-1660 Sequence 1660, Ap
3	234	15.0	1388	6	US-10-821-234-1143 Sequence 1143, Ap
4	226	14.5	795	6	US-10-770-726-49 Sequence 49, Appl
5	218.5	14.0	1404	6	US-10-878-556A-169 Sequence 169, Appl
6	212.5	13.6	718	6	US-10-878-556A-97 Sequence 97, Appl
7	209.5	13.4	1410	6	US-10-821-234-1050 Sequence 1050, Ap
8	193.5	12.4	1189	7	US-11-074-176-134 Sequence 134, Appl
9	188.5	12.1	531	7	US-11-060-914-4 Sequence 4, Appl1
10	186.5	11.9	1410	6	US-10-878-556A-136 Sequence 136, App
11	186.5	11.9	1586	6	US-10-821-234-901 Sequence 901, App
12	186	11.9	155	6	US-10-952-535A-17 Sequence 17, Appl
13	186	11.9	284	6	US-10-821-234-1332 Sequence 1632, Ap
14	185	11.8	664	6	US-10-878-556A-44 Sequence 44, Appl
15	184.5	11.7	558	6	US-10-467-657-694 Sequence 694, App
16	183.5	11.7	248	6	US-10-878-556A-175 Sequence 175, App
17	183	11.6	1565	6	US-10-467-657-2704 Sequence 2704, Ap
18	181	11.6	145	6	US-10-952-535A-19 Sequence 19, Appl
19	180.5	11.5	672	7	US-11-000-463-455 Sequence 455, App
20	179	11.5	121	6	US-10-952-535A-14 Sequence 14, Appl
21	178.5	11.4	1299	6	US-10-821-234-1145 Sequence 1145, App
22	176.5	11.3	1043	6	US-10-821-234-1055 Sequence 1055, App
23	175	11.2	989	6	US-10-821-234-975 Sequence 975, App
24	174	11.1	691	6	US-10-131-826A-16 Sequence 16, Appl
25	174	11.1	703	6	US-10-821-234-963 Sequence 963, App

26	173.5	11.1	898	7	US-11-099-691-7 Sequence 7, Appli
27	170	10.9	1122	6	US-10-467-657-6112 Sequence 6112, Ap
28	168	10.7	676	7	US-11-135-855-28 Sequence 28, Appl
29	168	10.7	717	7	US-11-135-855-29 Sequence 29, Appl
30	165.5	10.6	465	6	US-10-878-556A-116 Sequence 116, App
31	160.5	10.3	495	6	US-10-821-234-1154 Sequence 1154, App
32	158	10.1	600	6	US-10-878-556A-43 Sequence 43, Appl
33	157.5	10.1	1616	6	US-10-821-234-1497 Sequence 1497, Ap
34	155	9.9	585	6	US-10-878-556A-42 Sequence 42, Appl
35	154.5	9.9	418	6	US-10-878-556A-78 Sequence 78, Appl
36	150.5	9.6	459	6	US-10-821-234-896 Sequence 896, App
37	149	9.5	882	7	US-11-012-762-34 Sequence 34, Appl
38	148	9.5	1618	6	US-10-984-645-2 Sequence 2, Appl1
39	146	9.3	440	7	US-11-108-172-1059 Sequence 1059, App
40	146	9.3	3674	7	US-11-000-463-454 Sequence 454, App
41	145	9.3	1095	6	US-10-793-626-3154 Sequence 3154, Ap
42	144	9.2	1596	7	US-11-060-005-4 Sequence 4, Appl1
43	143.5	9.2	431	7	US-11-074-176-132 Sequence 132, App
44	143.5	9.2	639	6	US-10-821-234-907 Sequence 907, App
45	143	9.1	747	6	US-10-821-234-1662 Sequence 1662, Ap

ALIGNMENTS

```
RESULT 1
US-11-060-914-2
; Sequence 2, Application US/11060914
; Publication No. US20050261184A1
; GENERAL INFORMATION:
; APPLICANT: Kaufman, Paul L
; APPLICANT: Geiger, Benjamin
; APPLICANT: Bershidsky, Alexander
; APPLICANT: Borraas, Teresa
; TITLE OF INVENTION: Method for Treating Glaucoma
; FILE REFERENCE: 960296.00128
; CURRENT APPLICATION NUMBER: US/11/060,914
; CURRENT FILING DATE: 2005-02-18
; PRIOR APPLICATION NUMBER: US 60/545,722
; PRIOR FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: US 60/545,723
; PRIOR FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 2
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-060-914-2

Query Match      15.4%; Score 241; DB 7; Length 793;
Best Local Similarity 25.4%; Pred. No. 6.6e-06;
Matches 85; Conservative 73; Mismatches 105; Indels 72; Gaps 12;

4 EQERRAKEKLOEQ-----SDLEQDLRAKEKLOEQOSD-----LEQERRAKE 45
  :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
182 ENKEDKEKEKEEKEEKEPRGSGIGENOVEVWEKTTTSQETTVMSLKNQISSEBPQOE 241
  :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
46 KLOEQOSD-----LEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQDLRAKEKLOE 100
  :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
242 BERQGSSEISHHKMEEDKERAEARAEARERAEER--ERIKAQDDKIADERARIEAE 299
  :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
101 QOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQER 160
  :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
300 KAAQOEERRRAE---ERRRMRREEKRAAE---ERRRIKEERKRAE---ERRRIKEERK 350
  :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
161 RAKKLEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKTKYNER 220
  :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
351 PAAE---ERQIKKEERKRAE---ERQRRABEEERKAKYEEQKQKOLEK---KHAMOE 401
  :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
221 KKEHGDIILAEVLYRLLEIPALEPSENERGYIIPHOSLPODNRGNSRDSKEISIEKTN 280
  :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
402 TKIGEKVEQKIEBKWV-----NEKKAQEDLQTAVALKKQ 436
  :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

```

Oy      281 RESITTNVEGRDIIHGHLBEKKDGSIAPEQKEDK 315
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      437 GEEKSTKVQAKRE---KLQEDKPTFKKEIIDEK 467

```

RESULT 2

```

US-10-793-626-1660
; Sequence 1660, Application US/10793626
; Publication No. US20050255478A1
;
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU34800S
; CURRENT APPLICATION NUMBER: US/10/793, 626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1660
; LENGTH: 885
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; US-10-793-626-1660

```

Query Match	15.0%	Score 235;	DB 6;	Length 885;
Best Local Similarity	25.2%	Pred. No. 1.6e-07;		
Matches 89; Conservative	86;	Mismatches 90;	Indels 88;	Gaps 17.

Qy	4	EOBRRAKEKLOEOQ-----QSDLEODRLAKKELOEOQSDLEOBRRAKEKLOEOQS	52
Db	501	EOKTBA--EOLSEQYFASQOYNNLKQVNHHELDRLKTOEHLNNEH---EEFEFEKN	556
Qy	53	DLEOBRRAKEKLOEOQSDLEOBRRAKEKLOEOQSDLEQ--DRLAKE-----KLOEO	101
Db	557	DQYQSDSKSEKTELKEKNHLE---IQOOLKQSDLEERTYQULSKEGKASTHQTQOOLHOK	613
Qy	102	QSDLE-----EOBRRAKEKLOEOQSDLEOER--RAKEKLOEOQSDLEOBRRAKEKLOEOQ	153
Db	614	QSDDLAVVKERISQKQYVERLDBQQLSDSERQKLEVENEKIKLFPNSDMMGDABEKLKEQ--	672
Qy	154	SDLEOBRRAKEKLOEOQSDLEOER--AKEKLOEOQSDLE-----OBRRAKE--	198
Db	673	--IQOQENVRQNNQQLSEIKQQRKQKNEKLEINESQLOQCHQDILSIENHYQDIKAKQ	730
Qy	199	-----KLOEOQRDLQOKADTKKULERKKENGDILAEDLYRLLEIPAI	241
Db	731	KLDVLINHAIDHLNDTYQULVEARAMEYSDETIDILRKVKVLTXTMTIDEL--QPVNINAI	789
Qy	242	ELPSE--NEAGYIYPHOSSLPQDRNGSRDSKEI--SIIEKTNSESITTNIEGR	291
Db	790	EQFEETNEEYTFLEQO-----RTULREKKEKLEQOIHENMDE---VEGR	830

RESULT 3

```

US10-021-234-1143
/ Sequence 1143, Application US/10821234
/ Publication No. US20050255114A1
/ GENERAL INFORMATION:
/ APPLICANT: Labat, Ivan
/ APPLICANT: Stache-Crain, Birgit
/ APPLICANT: Andermann, Susan
/ APPLICANT: Tang, Y. Tom
/ TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
/ FILE REFERENCE: 821A
/ CURRENT APPLICATION NUMBER: US/10/821.234
/ CURRENT FILING DATE: 2004-04-07
/ PRIOR APPLICATION NUMBER: US 60/462,047
/ PRIOR FILING DATE: 2003-04-07
/

```

```

; NUMBER OF SEQ ID NOS: 1704
;
; SOFTWARE: pc_seq_genes Version 1.0
;
; SEQ ID NO 1143
;
; LENGTH: 1388
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
;
US-10-821-234-1143

```

Query Match	15.0%;	Score 234;	DB 6;	Length 1388;
Best Local Similarity	23.4%;	Pred. No. 2.8e-07;		
Matches	92;	Conservative	86;	Mismatches 98;
				Indels 118;
				Gaps 17;

[illegible]

RESULT 4

```

00000000 Sequence 49, Application US/107707026
00000001 Publication No. US200505266409A1
00000002 GENERAL INFORMATION:
00000003
00000004 APPLICANT: Wyeth
00000005 APPLICANT: Brown, Eugene
00000006 APPLICANT: Liu, Wei
00000007 TITLE OF INVENTION: COMPOSITIONS AND
00000008 FILE REFERENCE: AM101079 (031896-01)
00000009 CURRENT APPLICATION NUMBER: US/10/7707026
00000010 CURRENT FILING DATE: 2004-02-04
00000011 NUMBER OF SEQ ID NOS: 48640
00000012 SOFTWARE: PatentIn Version 3.2
00000013 SEQ ID NO 49
00000014 LENGTH: 795
00000015 TYPE: PRT
00000016 ORGANISM: Homo sapiens
00000017
00000018 US-10-770-726-49

```

Query Match	14.5%;	Score 226;	DB 6;	Length 795;
-------------	--------	------------	-------	-------------

Matches	76;	Conservative	78;	Mismatches	105;	Indels	78;	Gaps	12
---------	-----	--------------	-----	------------	------	--------	-----	------	----

[illegible]


```
Db      1092 LEEERERREERERKLGQWRQTOSLNPAPSPPLTAQOMKEKSTLQRPETVIRE----- 1146
Qy      46 KLQEQOSDLEQERRAKEKLQEQOSDLEQ-----ERRAKEKLQEQOSDLEQORLAK 96
Db      1147 -LQPOQOPPTERRDLQYTVTSKEELSSGTVCPPBPWKDAEKLEKQOMHIVMLKE 1205
Qy      97 KLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSD 156
Db      1206 -LQEQOSKPRKAEBSDRKLMLQEQOKLQEQOSKQEDDEBEEDDDVDVTLMLQR-L 1263
Qy      157 EOERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQORRAD 213
Db      1264 EAEERRARTMPAIVSLVDLQDEERRRQQLQEQEMRK-REAEADARQEQ-ERRRQEEER--- 1318
Qy      214 TKKLERKKEHGDILAEIDYGRLEIPALPSENERGYTPH-----QSSLPQD 262
Db      1319 TKRDAEKKRRQ-----EEGYSRLEAEERRRQHDPAARLLEPAPGLCRPPLPRD 1368
```

```
RESULT 8
US-11-074-176-134
; Sequence 134, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Ruseell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McCallife, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; FILE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-134
```

```
Query Match      12.4%; Score 193.5; DB 7; Length 1189;
Best Local Similarity 21.0%; Pred. No. 5.1e-05;
Matches 74; Conservative 89; Mismatches 105; Indels 85; Gaps 14;
```

```
Qy      4 EOERRAKEKLQEQOSDLEQ-----EODRLAKEKLQEQOSDLEQERRAK- 44
Db      172 QOKEEQONOLKQTDVLRINDLVKELSERLPELNEQSSLAKE-YFQKSGDLKXKSL 230
Qy      45 -----EKLOEQOSDLEQ-----BRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQOR 92
Db      231 APEIENINQORREDIQKSAQNKILLAKLQDEVDQSAATYQKAEYKQLRDER-DHTQNK 289
Qy      93 LAK--EKLOEQOSDLEQ-----EOERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKE 147
Db      290 LKLSKDLSEINSLQMAEQSRQFDDATKEEYKN--QVQKQONLQVLKADLDLKLKEXK 347
Qy      148 KLQEQOSDLEQER-----RAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLE 191
Db      348 KLQEDQDVAKTEKGQLTGEINEDPEELNKLDIDIRNNWQQLQDQTTNNQVNLNSDL- 406
Qy      192 QERRAKEKLQEQOSDLEQORRADTKKLERKKEHGDILAEIDYGRLEIPALPSENERGY 251
Db      407 ---RRSQADTTYQIGDVSQKLTDAQKLE-----QLAIEGKLTDRKQK--- 447
Qy      252 YIPQSSLPQDNGNSRDSKEISIT-----EKTNRESITTVNVEGRDILHKGH 298
Db      448 ---EQNAIVIRKQNNQNLTELNLNQVVAERNELEKVARREALVNIQKRH 497
```

```
RESULT 9
US-11-060-914-4
; Sequence 4, Application US/11060914
; Publication No. US20050261184A1
; GENERAL INFORMATION:
; APPLICANT: Kaufman, Paul L
; APPLICANT: Geiger, Benjamin
; APPLICANT: Borshady, Alexander
; APPLICANT: Borras, Teresa
; TITLE OF INVENTION: Method for Treating Glaucoma
; FILE REFERENCE: 960296.00128
; CURRENT APPLICATION NUMBER: US/11/060,914
; CURRENT FILING DATE: 2005-02-18
; PRIOR APPLICATION NUMBER: US 60/545,722
; PRIOR FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: US 60/545,723
; PRIOR FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-11-060-914-4
```

```
Query Match      12.1%; Score 188.5; DB 7; Length 531;
Best Local Similarity 22.1%; Pred. No. 4.7e-05;
Matches 77; Conservative 82; Mismatches 136; Indels 53; Gaps 12;
```

```
Qy      5 QERRAK-----EKLOEQOSDLEODRLAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQOR 58
Db      98 EERRQKRLQEALERQKEDPTTDSLSVPSRRMNNNAENETABEGEESGRGYEME 157
Qy      59 RAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQORLAKKEKLQEQOSDLEQERRAKEKLQEQO 118
Db      158 ETEVAVITSYQKNSYQDABDKK-EKEKEEKEEKUKGNLGENQIKDEKIKQKPEKEV 216
Qy      119 QSDLEQERRAKE-----KLQEQOSDLEQER-----RAKE 147
Db      217 KNFLDRKKGFTVEVKAQNGEFMTKQKOTENAFSPRSGRASGDKEBAGAPQVNAKRL 276
Qy      148 KLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQ-----BRAKEKLQEQOR 205
Db      277 ELRRRGETESEEFKLLQKQOEALTELE-ELKKRBERRKVLEEBEQRRKQOEADRKAR 335
Qy      206 DLQKADTKKLERKKEHGDILAEIDYGRLEIPALPSENERGY--YIPQSSLPQD 263
Db      336 E-EERKRRLKEIERRR-----AEAAEKQKMPEDGL-SEDKKPKCFTPKGSSLKIBE 387
Qy      264 RGN-SRDSKEISITIEKTNRESITTVNVEGRDILHKGHL-EKKDGSIKP 309
Db      388 RAELFKNSVQKSGVASTHOAAVVSKIDSRLEQYTNALBETKASKMPKP 435
```

```
RESULT 10
US-10-878-556A-136
; Sequence 136, Application US/10878556A
; Publication No. US2005026399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 136
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/trbl_human
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:12:18 ; Search time 8.91686 Seconds

(without alignments)
1262.481 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603

Sequence: 1 LOEQORDLEQKADTKKLE.....GHLEKKDGSIKPKEDKS 117

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	602	99.8	1909	2	A45592
2	102	16.9	210	2	C70309
3	94.5	15.7	772	2	I50463
4	91.5	15.2	635	1	I37060
5	90.5	15.0	1085	2	S55352
6	90	14.9	756	2	T00367
7	89.5	14.8	3488	2	T34418
8	88	14.6	762	2	G88436
9	88	14.6	791	2	T24435
10	87	14.4	550	2	T32560
11	87	14.4	779	2	E54024
12	84.5	14.0	585	1	A24168
13	84.5	14.0	605	1	I37061
14	84.5	14.0	765	1	ISHUT1
15	84	13.9	913	2	T52485
16	83.5	13.8	615	2	T06108
17	83.5	13.8	789	2	T51310
18	83.5	13.8	821	2	S67087
19	83.5	13.8	1025	2	A47453
20	83.5	13.8	1706	2	T39305
21	83	13.8	483	2	S41853
22	83	13.8	684	2	S52835
23	82.5	13.7	495	1	I37062
24	82.5	13.7	1027	2	S37711
25	81.5	13.5	442	2	D70315
26	81.5	13.5	980	2	E71606
27	81.5	13.5	1829	2	T26135
28	81	13.4	543	2	T32973
29	81	13.4	1390	2	T14004

30	80.5	13.3	678	2	A54514	glutamic acid-rich
31	80.5	13.3	768	2	H54024	protein kinase (EC
32	80.5	13.3	777	2	B54024	protein kinase (EC
33	80.5	13.3	1032	2	I38510	neuronal kinesin h
34	80.5	13.3	1804	2	T34518	nestin - golden ha
35	80	13.3	452	2	D64583	hypothetical prote
36	80	13.3	729	2	T50989	hypothetical prote
37	79.5	13.2	501	2	C71948	hypothetical prote
38	79.5	13.2	816	2	S64439	hypothetical prote
39	79.5	13.2	1132	2	T43483	translational initia
40	79.5	13.2	1203	2	S26650	DNA-binding protei
41	79.5	13.2	1840	2	T29091	translating transcrip
42	79.5	13.1	511	2	H83339	probable transcrip
43	79	13.1	630	2	S29796	hypothetical prote
44	79	13.1	721	2	S29795	hypothetical prote
45	79	13.1	783	2	A55817	cyclin-dependent k

ALIGNMENTS

RESULT 1
A45592
liver stage antigen ISA-1 - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 22-Nov-1993 #sequence revision 02-Dec-1994 #text change 09-Jul-2004
C:Accession: S24597; A45592; S29393; S34842; B45592; C45592; D45592
R:Zhu, J.; Hollingdale, M.
Submitted to the EMBL Data Library, November 1990
A:Reference number: S24597
A:Accession: S24597
A:Molecule type: DNA
A:Residues: 1-1909 <ZHU>
A:Cross-references: UNIPROT:Q25893; UNIPARC:UPI0000081753; EMBL:X56203; NID:99915; PID:
R:Zhu, J.; Hollingdale, M.R.
Mol. Biochem. Parasitol. 48, 223-226, 1991
A:Title: Structure of Plasmodium falciparum liver stage antigen-1.
A:Reference number: A45592; MUID:92107224; PMID:1840628
A:Accession: A45592
A:Molecule type: DNA
A:Residues: 1-195,638-688,1165-1215,1590-1909 <ZH2>
A:Cross-references: UNIPARC:UPI0000177378; UNIPARC:UPI0000177379; UNIPARC:UPI000017737F
A:Note: Sequence extracted from NCBI backbone (NCBI:75012, NCBI:75014, NCBI:75015)
A:Title: A liver-stage-specific antigen of Plasmodium falciparum characterized by gene
A:Reference number: S29393; MUID:87315391; PMID:3306406
A:Accession: S29393
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 323-387 <GUE1>
A:Cross-references: UNIPARC:UPI000017737C; EMBL:M28266
R:Guerin-Marchand, C.; Druilhe, P.; Galey, B.; Londono, A.; Patapoutikou, J.; Beaudou
Submitted to the EMBL Data Library, April 1992
A:Description: A liver-stage-specific antigen of plasmodium falciparum characterized by
A:Reference number: S34842
A:Accession: S34842
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 323-381, 'HKA1', 'GUE2'
A:Cross-references: UNIPARC:UPI000017737D; EMBL:M28266
A:Note: difference at carboxyl end due to frameshift error
C:Comment: This protein is found as flocculent material in the parasitophorous vacuole.
C:Superfamily: trichohyalin; calmodulin repeat homology
C:Keywords: EF hand
F:154-1629/Region: 17-residue repeats (A-K-E-K-L-Q-E-Q-Q-S-D-L-E-Q-E-R-R-R)
Query Match 99.8% Score 602; DB 2; Length 1909;
Best Local Similarity 99.1% Pred. No. 1.1e-43;
Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
1 LOEQORDLEQKADTKKLEGGDILADLYGRLETPAIEPSENERGYVPHOSSL 60
|||||

```
Db      1620 LQEOORDEQRKADTKNLERKKEHGDVLAEDLYGRLEIPALIELPSENERGYIIPHOSL 1679
Qy      61 PDQNGNSRDSKEISITTEKNTRESITTVNTEGRDRIHKGHEKKGSIKPEQEKDS 117
Db      1680 PDQNGNSRDSKEISITTEKNTRESITTVNTEGRDRIHKGHEKKGSIKPEQEKDS 1736

RESULT 2
C70309
transcription regulator - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: C70309
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oy
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: C70309
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-210 <NOF>
A:Cross-references: UNIPROT:O6502; UNIPARC:UPI000056271; GB:AE000673; NID:92982834; PI
A:Experimental source: strain VF5
C:Genetics:
A:Gene: htn

Query Match      16.9%; Score 102; DB 2; Length 210;
Best Local Similarity 28.9%; Pred. No. 0.097;
Matches 43; Conservative 20; Mismatches 40; Indels 46; Gaps 7;

Qy      3 EQQRDLQQRKAD-----PKNLERKKEHGDVLAEDLYGRLEIPALIELPSENERGYIIPH 56
Db      33 ERIGDLPRKADVRIITAIATNKNLKLVEBSK-FREDLYRLANVNVLEPPLERKEDITH 91
Qy      57 QSS-----LPQNGNSRDSKEISITTEKNTRESI---TTN 88
Db      92 LVSYLEFKSKHGKVKGVSPAMKLLLEVDYFGVRE-----LNNIAHAVVVSSTS 145
Qy      89 VEGRADIHGKHEKKDGS-TPQEQEKD 116
Db      146 LIGPDDL-----EIRIGSKVKKKPRENE 170

RESULT 3
I50463
protein kinase - chicken
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 05-Oct-2004
C:Accession: I50463
R:Li, H.; Grenet, J.; Valentin, M.; Lahti, J.M.; Kidd, V.J.
Gene 153, 237-242, 1995
A:Title: Structure and expression of chicken protein kinase PITSLRE-encoding genes.
A:Reference number: I50463; MUID:95180725; PMID:7875595
A:Accession: I50463
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-772 <PIX>
A:Cross-references: UNIPROT:Q91013; UNIPARC:UPI00000FC347; EMBL:U16656; NID:9571458; PID
C:Genetics:
A:Gene: cPITSLRE
A:Introns: 37/3; 76/2; 118/1; 164/2; 208/1; 249/1; 283/1; 329/1; 351/1; 407/3; 438/1; 47
C:Keywords: ATP
F:413-668/Domain: protein kinase homology <KIN>
F:421-429/Region: protein kinase ATP-binding motif

Query Match      15.7%; Score 94.5; DB 2; Length 772;
Best Local Similarity 29.3%; Pred. No. 1.9;
Matches 36; Conservative 20; Mismatches 48; Indels 19; Gaps 6;

Qy      1 LQEQORDLEQ-RKADTK--KNLERKKEHGDVLAEDLYGRLEIPALIELPSENERGYIIPHQ 57
Db      17 LQKKRKKEDEKAEIKGMNSDDRDSKDSLE---GEIARDHMETTINS-----PYR 68
```

```
Qy      58 SLPQDNGNSRDS-----KEISITTEKNTRESITTVNTEGRDRIHKGHEKKGSIKPE 111
Db      69 REDSMEDGGEEDSLAIRPQOMSREKETHHRKDEKREKR--HRSASAEKHAHVXEK 126
Qy      112 QKE 114
Db      127 ERE 129

RESULT 4
I137060
involucrin L - gorilla
C:Species: Gorilla gorilla (gorilla)
C:Date: 16-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: I137060
R:Reuter, J.; Green, H.
Proc. Natl. Acad. Sci. U.S.A. 86, 1283-1286, 1989
A:Title: Divergent evolution of part of the involucrin gene in the hominoids: Unique int
A:Reference number: I137060; MUID:89145226; PMID:2919176
A:Accession: I137060
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-635 <RES>
A:Cross-references: UNIPROT:Q28423; UNIPARC:UPI0000087628; GB:M23603; GB:J04499; NID:934
C:Comment: During the terminal differentiation of keratinocytes, this protein from the c
linked envelope under the plasma membrane.
C:Superfamily: involucrin
C:Keywords: cornified cell envelope; duplication; epidermis; tandem repeat
F:153-590/Region: 10-residue repeats (Q-E-G-Q-[PLV]-[KE]-[LH]-[PL]-E-Q)

Query Match      15.2%; Score 91.5; DB 1; Length 635;
Best Local Similarity 26.1%; Pred. No. 2.8;
Matches 36; Conservative 29; Mismatches 42; Indels 31; Gaps 8;

Qy      2 QEQQRDLQQRKADTKNLERKKEHGDVLAEDLYGRLEIP-----AIELP 45
Db      149 KQQLLEPQQRGELKHLQEQK--GQLLEPQEQEGQLLEPQEQEGQLKHLQEQKQQLLEP 206
Qy      46 SENERGYIIPHS-----SLPQDNGNSR--DSKEISITTEKNTRESITTVNTEGRDRIHKG 99
Db      207 EQEQQLLEPQEQEGQLLEPQEQEGQLKHLQEQEGQLLEPQEQEGQLKHLQEQ--KQ 262
Qy      100 LE-EKKDGSIR-PEKE 114
Db      263 LELPQEQEGQLLEPQEQ 280

RESULT 5
S55352
IFI1 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein L8083.9; protein YLR223c; RRP3 protein
C:Species: Saccharomyces cerevisiae
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S55352; S51446; S47477
R:Cherel, I.; Thuriaux, P.
Yeast 11, 261-270, 1995
A:Title: The IFI1 gene product interacts with a fork head protein in Saccharomyces cere
A:Reference number: S55352; MUID:95304839; PMID:7785326
A:Accession: S55352
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1085 <CHE>
A:Cross-references: UNIPROT:P39520; UNIPARC:UPI000012D3D6; EMBL:Z29488; NID:9531491; PI
R:Hallsworth, K.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmid 8083.
A:Reference number: S51443
A:Accession: S51446
A:Molecule type: DNA
A:Residues: 1-1085 <HAL>
A:Cross-references: UNIPARC:UPI000012D3D6; EMBL:U19027; NID:9609363; PID:9609372; MIPS:
```


[illegible]

8; Teumer, J.; Green, H.

Proc. Natl. Acad. Sci. U.S.A. 86, 1283-1286, 1989

A>Title: Divergent evolution of part of the involucrin gene in the hominoids: Unique int

A/Reference number: 137060; PMID:89145226; PMID:2919176

A/Accession: 137061

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-605 <RES>

A/Cross-references: UNIPROT:Q28424; UNIPARC:UPI000008710D; GB:M23604; GB:J04499; NID:934

C/Comment: During the terminal differentiation of keratinocytes, this protein from the c

C/Superfamily: Involucrin

C/Keyword: cornified cell envelope; duplication; epidermis; tandem repeat

F:153-561/Region: 10-residue repeats (Q-E-G-Q-[PLV]-[KE]-[LH]-[PL]-E-Q)

Query Match 14.0%; Score 84.5; DB 1; Length 605;

Best Local Similarity 23.5%; Pred. No. 11;

Matches 32; Conservative 25; Mismatches 32; Indels 47; Gaps 7;

QY 2 QEQORDLEQRKATYKTKLKKKEHGD-LAEDLYGRLEIF-----AIELP 45

DB 149 KEQLLEPEQEGQKLEQOK--GQLEPEQEGQLEPEQEGQKLEQOKGQLELP 206

QY 46 SENERGYIYPHQS-----SLPDNRGNSRDSKEISIEKTNRESITTVNVEGRDINKHGLE 101

DB 207 EQEGQLEPEQEGQLEPEQEGQLEPEQEGQKLEQ-----QEQGLE 244

QY 102 -EKQDGSIK-PEQKE 114

DB 245 LPEQEGQLEPEQEGQ 260

RESULT 14

ISHUT1

DNA topoisomerase (EC 5.99.1.2) - human

N/Alternate names: nicking-closing enzyme; omega-protein; relaxing enzyme; swivelase; ty

C/Species: Homo sapiens (man)

C/Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004

C/Accession: A30887; A40008; S13821; S02397; S40643; A34422

R/D/Arpa, P.; Machlin, P.S.; Reiter, H.; Reichfield, N.F.; Cleveland, D.W.; Earnshaw,

Proc. Natl. Acad. Sci. U.S.A. 85, 2543-2547, 1988

A>Title: CDNA cloning of human DNA topoisomerase I: catalytic activity of a 67.7-kDa cat

A/Reference number: A30887; PMID:88190108; PMID:2833744

A/Accession: A30887

A/Molecule type: mRNA

A/Residues: 1-765 <DAR>

A/Cross-references: UNIPROT:P11387; UNIPARC:UPI00001421A9; GB:J03250; NID:G339805; PIDN:

R/Kunze, N.; Yang, G.; Doelberg, M.; Sundarp, R.; Knippers, R.; Richter, A.

J. Biol. Chem. 266, 9610-9616, 1991

A>Title: Structure of the human type I DNA topoisomerase gene.

A/Reference number: A40008; PMID:91236733; PMID:1851751

A/Accession: A40008

A/Molecule type: DNA

A/Residues: 1-144, 'A', 146-553, 'E', 555-765 <KUN>

A/Cross-references: UNIPARC:UPI00001730AD; GB:M60688; GB:M60689; GB:M60690; GB:M60691; C

M60702; GB:M60703; GB:M60704; GB:M60705; GB:M60706

R/Kunze, N.; Klein, M.; Richter, A.; Knippers, R.

Eur. J. Biochem. 194, 323-330, 1990

A>Title: Structural characterization of the human DNA topoisomerase I gene promoter.

A/Reference number: S13821; PMID:91099302; PMID:2176592

A/Accession: S13821

A/Molecule type: DNA

A/Residues: 1-20 <KUN>

A/Cross-references: UNIPARC:UPI00001730AE; EMBL:X53601

R/Oddou, P.; Schmid, U.; Knippers, R.; Richter, A.

Eur. J. Biochem. 177, 523-529, 1988

A>Title: Monoclonal antibodies neutralizing mammalian DNA topoisomerase I activity.

A/Reference number: S02397; PMID:89064806; PMID:2461859

A/Accession: S02397

A/Molecule type: mRNA

A/Residues: 344-765 <ODD>

A/Cross-references: UNIPARC:UPI00001730AF; GB:M60657

R/Tamura, H.O.; Kohchi, C.; Yamada, R.; Ikeda, T.; Koiwai, O.; Patterson, E.; Keene, J.F

Nucleic Acids Res. 19, 69-75, 1991

A>Title: Molecular cloning of a cDNA of a camptothecin-resistant human DNA topoisomerases

A/Reference number: S40643; PMID:91187651; PMID:1849260

A/Accession: S40643

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 523-543; 573-582, 'D', 584-593 <TAU>

A/Cross-references: UNIPARC:UPI00001730B0; UNIPARC:UPI00001730B1

R/Mul, G.G.; Jimenez, S.A.; Riggs, E.; Ziemnicka-Kocula, D.

Proc. Natl. Acad. Sci. U.S.A. 86, 8492-8496, 1989

A>Title: Determination of an epitope of the diffuse systemic sclerosis marker antigen D

ity in systemic sclerosis.

A/Reference number: A34422; PMID:90046823; PMID:2479024

A/Accession: A34422

A/Molecule type: mRNA

A/Residues: 657-765 <MAU>

A/Cross-references: UNIPARC:UPI000016B0FE; GB:M27913; NID:G339807; PIDN:AAA61208.1; PID

C/Comment: Type I DNA topoisomerase catalyzes the ATP-independent transient breakage of

in another, followed by rejoining. This reaction will lead to the conversion of one to

C/Genetics:

A/Genes: GDB:TOP1

A/Cross-references: GDB:120444; OMIM:126420

A/Map position: 20q12-20q13.1

A/Introns: 11/3; 20/1

A/Note: the list of introns is incomplete

C/Superfamily: eukaryotic type I DNA topoisomerase

C/Keyword: DNA binding; DNA replication; heterotetramer; isomerase

F:723/Active site: Tyr #status predicted

Query Match 14.0%; Score 84.5; DB 1; Length 765;

Best Local Similarity 25.6%; Pred. No. 14;

Matches 31; Conservative 22; Mismatches 47; Indels 21; Gaps 5;

QY 3 EQQRDLQRKATYKTKLKKKEHGD-LAEDLYGRLEIFALPSENERGYIYPHQSL 60

DB 75 EKHKDKQKQKDKKREKRAAGDQKIKKEKNGSSPQIDPEDD-GYFVPPKEDI 133

QY 61 -----PDNRGNSRDSKEISIEKTNRESITTVNVEGRDINKHGLEKXKXGSIKPEQED 115

DB 134 KPLKRPDEDDVDYKPKIK-TEDTYKKE-----KKLEEEEDGKLKKPKND 180

QY 116 K 116

DB 181 K 181

RESULT 15

T52485

neurofilament protein NF-M(2), middle molecular weight (imported) - African clawed frog

C/Species: Xenopus laevis (African clawed frog)

C/Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004

C/Accession: T52485

R/Gervasi, C.; Szabo, B.G.

Brain Res. Mol. Brain Res. 48, 229-242, 1997

A>Title: Sequence and expression pattern of two forms of the middle molecular weight ne

A/Reference number: Z26090

A/Accession: T52485

A/Status: preliminary; translated from GB/EMBL/DDB1

A/Molecule type: mRNA

A/Residues: 1-913 <GER>

A/Cross-references: UNIPROT:O13099; UNIPARC:UPI00000FCFAD; EMBL:U85970; PIDN:AAB53390.1

Query Match 13.9%; Score 84; DB 2; Length 913;

Best Local Similarity 26.5%; Pred. No. 19;

Matches 30; Conservative 21; Mismatches 50; Indels 12; Gaps 4;

QY 2 QEQORDLEQRKATYKTKLKKKEHGD-LAEDLYGRLEIFALPSENERGYIYPHQSLP 61

DB 550 EEEEGGEGEEYVKAEEEEKKEGEGDAE---VEETKAEDEEGEKG-----EEBAE 600

QY 62 QDNRGNSRDSKEISIEKTNRESITTVNVEGRDINKHGLEKXKXGSIKPEQKE 114

DB 601 KEKGEEEEKAEVAABE-GEEGEEEGEGEKEDKG--EEKEGEAAEAAE 650

Search completed: December 7, 2005, 07:21:23
Job time : 10.9169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:11:08 ; Search time 37.0165 Seconds
(without alignments)
2229.881 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LQEQQRDLQRKADTKKNL.....GHLEKKDGSIKPEQKEDKS 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot_05.80.*
1: uniprot_05.80.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	603	100.0	493	Q25886_PLAFA	Q25886 plasmodium
2	602	99.8	1909	Q25893_PLAFA	Q25893 plasmodium
3	597	99.0	1596	Q81J44_PLAF7	Q81J44 plasmodium
4	554	91.9	280	Q25844_PLAFA	Q25844 plasmodium
5	554	91.9	280	Q25849_PLAFA	Q25849 plasmodium
6	554	91.9	280	Q25850_PLAFA	Q25850 plasmodium
7	553	91.7	280	Q25843_PLAFA	Q25843 plasmodium
8	553	91.7	280	Q25846_PLAFA	Q25846 plasmodium
9	553	91.7	280	Q25847_PLAFA	Q25847 plasmodium
10	553	91.7	280	Q25851_PLAFA	Q25851 plasmodium
11	550	91.2	280	Q25848_PLAFA	Q25848 plasmodium
12	550	91.2	280	Q25852_PLAFA	Q25852 plasmodium
13	549	91.0	280	Q25845_PLAFA	Q25845 plasmodium
14	549	91.0	280	Q25853_PLAFA	Q25853 plasmodium
15	549	91.0	280	Q25854_PLAFA	Q25854 plasmodium
16	549	91.0	280	Q25855_PLAFA	Q25855 plasmodium
17	549	91.0	280	Q25888_PLAFA	Q25888 plasmodium
18	549	91.0	280	Q25889_PLAFA	Q25889 plasmodium
19	549	91.0	280	Q25900_PLAFA	Q25900 plasmodium
20	549	91.0	280	Q27243_PLAFA	Q27243 plasmodium
21	469	77.8	264	Q9GTX5_PLAFA	Q9GTX5 plasmodium
22	224	37.1	42	Q8WZK9_PLAFA	Q8WZK9 plasmodium
23	102	16.9	210	Q66502_AQUB	Q66502 equifex ae
24	98.5	16.3	924	Q81445_PLAF7	Q81445 plasmodium
25	97	16.1	1130	Q59YV6_CANAL	Q59YV6 candida alb
26	94.5	15.7	772	Q91013_CHICK	Q91013 gallus galli
27	94	15.6	245	Q51017_ENTHI	Q51017 entamoeba h
28	94	15.6	1813	Q51393_ENTHI	Q51393 entamoeba h
29	93.5	15.5	3812	Q4RLC8_TERNI	Q4RLC8 tetraodon n
30	92	15.3	603	Q9AHN3_PASWU	Q9AHN3 pascuteurella
31	92	15.3	783	Q4T6N6_TERNI	Q4T6N6 tetraodon n

32	91.5	15.2	635	Q28423_9PRIM	Q28423 gorilla gor
33	91	15.1	294	Q895S6_CLORE	Q895S6 clostridium
34	90.5	15.0	1085	1 IFH1 YEAST	P39520 saccharomyc
35	90	14.9	225	Q7Z5E9_HUMAN	Q7Z5E9 homo sapien
36	90	14.9	756	1 RPIP3_HUMAN	Q75154 homo sapien
37	90	14.9	756	Q4VXV7_HUMAN	Q4VXV7 homo sapien
38	89.5	14.8	295	Q6BXX9_DEBIA	Q6BXX9 debaryomyc
39	89.5	14.8	485	Q75EU2_ASHO	Q75EU2 ashya goss
40	89.5	14.8	18519	Q81SF7_CAEEL	Q81SF7 caenorhabdi
41	89.5	14.8	18519	Q81SF7_CAEEL	Q81SF7 caenorhabdi
42	88.5	14.7	752	Q4Z6S0_PLABE	Q4Z6S0 plasmodium
43	88.5	14.7	807	Q5TM60_MACMU	Q5TM60 macaca mula
44	88.5	14.7	847	Q4FPL4_9RICK	Q4FPL4 candidatus
45	88.5	14.7	1405	Q4Z3Q7_PLABE	Q4Z3Q7 plasmodium

ALIGNMENTS

```
RESULT 1
Q25886_PLAFA PRELIMINARY; PRT; 493 AA.
AC Q25886;
DT 01-NOV-1996 (TREMUREL. 01, Created)
DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)
DT 01-OCT-2002 (TREMUREL. 22, Last annotation update)
DE Liver stage antigen-1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5833;
RX MEDLINE=9426724; Pubmed=7515922;
RA Fidock D.A., Gras-Masse H., Lepers J., Brahimi K., Benmohamed L.,
RA Meilhou S., Langley G., Rousillon C., Tartar A., Druille P.;
RA "The plasmodium falciparum liver stage antigen LSA-1 is well conserved
RT and harbors major B- and T-cell epitopes.";
RL J. Immunol. 0:0-0(1994).
DR EMBL, Z30320; CAA82975.1; -; Genomic_DNA.
FT NON_TER
SQ SEQUENCE 493 AA; 59062 MW; FE252CC44F46996E CRC64;

Query Match 100.0%; Score 603; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 1.7e-41;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEQQRDLQRKADTKKNLEKKGHDILADLYGRLEIPALIEPSENGGYIPHOSSL 60
DB 204 LQEQQRDLQRKADTKKNLEKKGHDILADLYGRLEIPALIEPSENGGYIPHOSSL 263

QY 61 PDRKGRSRSKESITTEKNTRESITTNVGRPDHKGHEKKDGSIKPEQKEDKS 117
DB 264 PDRKGRSRSKESITTEKNTRESITTNVGRPDHKGHEKKDGSIKPEQKEDKS 320

RESULT 2
Q25893_PLAFA PRELIMINARY; PRT; 1909 AA.
AC Q25893;
DT 01-NOV-1996 (TREMUREL. 01, Created)
DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)
DT 01-JUN-2003 (TREMUREL. 24, Last annotation update)
DE Liver stage antigen.
```

GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NF54;
RC MEDLINE=92107224; PubMed=1840628; DOI=10.1016/0166-6851(91)90117-0;
RA Zhu J., Hollingdale M.R.; Plasmodium falciparum liver stage antigen-1.";
RT "Structure of Plasmodium falciparum liver stage antigen-1.";
RL Mol. Biochem. Parasitol. 48:223-226(1991).
DR EMBL; X56203; CAA39663.1; -; Genomic_DNA.
DR PIR; S24597; A45592.
DR HSSP; O15813; 1D7M.
SQ SEQUENCE 1909 AA; 230153 MW; 8B861A524C97217 CRC64;

Query Match 99.8%; Score 602; DB 2; Length 1909;
Best Local Similarity 99.1%; Pred. No. 9.7e-41;
Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALPEPSENERGYIIPHOSSL 60
DB 1620 LOEQORDLEQRKADTKNLERKKEHGDVLAEDLYGRLEIPALPEPSENERGYIIPHOSSL 1679

QY 61 PONDNRGNSRDSKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 117
DB 1680 PONDNRGNSRDSKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 1736

RESULT 3
Q81J44_PLAF7
ID Q81J44_PLAF7 PRELIMINARY; PRT: 1596 AA.
AC Q81J44;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Liver stage antigen, putative.
GN ORFNames=PF10.0356;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=22255705; PubMed=12368664; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung B., White O., Berriman M., Hyman R.W.,
RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow J., Suh B., Peterson J., Angiuoli S.,
RA Perera M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairhead A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Sudramanlan G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrett B.G.;
RT "Genome sequence of the human malaria parasite Plasmodium falciparum.";
RT Nature 419:498-511(2002).
DR EMBL; A6014834; AAN3553.1; -; Genomic_DNA.
SQ SEQUENCE 1596 AA; 191691 MW; COB3A2CC0F5D60AC CRC64;

Query Match 99.0%; Score 597; DB 2; Length 1596;
Best Local Similarity 98.3%; Pred. No. 2e-40;
Matches 115; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LOEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALPEPSENERGYIIPHOSSL 60
DB 1307 LOEQORDLEQRKADTKNLERKKEHGDVLAEDLYGRLEIPALPEPSENERGYIIPHOSSL 1366

QY 61 PONDNRGNSRDSKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 117
DB 1367 PONDNRGNSRDSKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 1423

RESULT 4

Q25844_PLAFA
ID Q25844_PLAFA PRELIMINARY; PRT: 280 AA.
AC Q25844;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates.";
RT Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40885; AAB59231.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32940 MW; E9708B3CFPAEA9CF CRC64;

Query Match 91.9%; Score 554; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 9.8e-38;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RRAADTKNLERKKEHGDILAEDLYGRLEIPALPEPSENERGYIIPHOSSLPONDNRGNSRD 70
DB 1 RRAADTKNLERKKEHGDILAEDLYGRLEIPALPEPSENERGYIIPHOSSLPONDNRGNSRD 60

QY 71 SKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 117
DB 61 SKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 107

RESULT 5
Q25849_PLAFA
ID Q25849_PLAFA PRELIMINARY; PRT: 280 AA.
AC Q25849;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates.";
RT Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40890; AAC41600.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32966 MW; E96D255154DEA9CF CRC64;

Query Match 91.9%; Score 554; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 9.8e-38;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RRAADTKNLERKKEHGDILAEDLYGRLEIPALPEPSENERGYIIPHOSSLPONDNRGNSRD 70
DB 1 RRAADTKNLERKKEHGDILAEDLYGRLEIPALPEPSENERGYIIPHOSSLPONDNRGNSRD 60

QY 71 SKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 117
DB 61 SKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 107

```
RESULT 6
Q25850_PLAFA
ID Q25850_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25850;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
  Hawley W.A., Collins W.E., Lal A.A.;
  "Sequence variations in the non-repetitive regions of the liver stage-
  RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
  RT isolates.";
  RL Mol. Biochem. Parasitol. 71:291-294(1995).
  EMBL; L40891; AAC41601.1; -; Genomic_DNA.
  FT NON TER 1
  SQ SEQUENCE 280 AA; 32989 MW; B96F812CFABEBD8 CRC64;

Query Match 91.7%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.2e-37;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 RKADTKKLERKKEHGDILAEADLYGRLEIPALPELPSENERGYIIPHOSSLPODNRGNSRD 70
DB 1 RKADTKKLERKKEHGDILAEADLYGRLEIPALPELPSENERGYIIPHOSSLPODNRGNSRD 60

QY 71 SKEISIIKTNRESITTNVEGRDIIKHGHLEKKDGSIKPEQKEDKS 117
DB 61 SKEISIIKTNRESITTNVEGRDIIKHGHLEKKDGSIKPEQKEDKS 107

RESULT 7
Q25843_PLAFA
ID Q25843_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25843;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
  Hawley W.A., Collins W.E., Lal A.A.;
  "Sequence variations in the non-repetitive regions of the liver stage-
  RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
  RT isolates.";
  RL Mol. Biochem. Parasitol. 71:291-294(1995).
  EMBL; L40884; AAB59230.1; -; Genomic_DNA.
  FT NON TER 1
  SQ SEQUENCE 280 AA; 32926 MW; 467080F3FBAEDJ3 CRC64;

Query Match 91.7%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.2e-37;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 RKADTKKLERKKEHGDILAEADLYGRLEIPALPELPSENERGYIIPHOSSLPODNRGNSRD 70
DB 1 RKADTKKLERKKEHGDILAEADLYGRLEIPALPELPSENERGYIIPHOSSLPODNRGNSRD 60
```

```
QY 71 SKEISIIKTNRESITTNVEGRDIIKHGHLEKKDGSIKPEQKEDKS 117
DB 61 SKEISIIKTNRESITTNVEGRDIIKHGHLEKKDGSIKPEQKEDKS 107

RESULT 8
Q25846_PLAFA
ID Q25846_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25846;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
  Hawley W.A., Collins W.E., Lal A.A.;
  "Sequence variations in the non-repetitive regions of the liver stage-
  RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
  RT isolates.";
  RL Mol. Biochem. Parasitol. 71:291-294(1995).
  EMBL; L40887; AAC41597.1; -; Genomic_DNA.
  FT NON TER 1
  SQ SEQUENCE 280 AA; 32882 MW; 46763641C81AFC33 CRC64;

Query Match 91.7%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.2e-37;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 RKADTKKLERKKEHGDILAEADLYGRLEIPALPELPSENERGYIIPHOSSLPODNRGNSRD 70
DB 1 RKADTKKLERKKEHGDILAEADLYGRLEIPALPELPSENERGYIIPHOSSLPODNRGNSRD 60

QY 71 SKEISIIKTNRESITTNVEGRDIIKHGHLEKKDGSIKPEQKEDKS 117
DB 61 SKEISIIKTNRESITTNVEGRDIIKHGHLEKKDGSIKPEQKEDKS 107

RESULT 9
Q25847_PLAFA
ID Q25847_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25847;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
  Hawley W.A., Collins W.E., Lal A.A.;
  "Sequence variations in the non-repetitive regions of the liver stage-
  RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
  RT isolates.";
  RL Mol. Biochem. Parasitol. 71:291-294(1995).
  EMBL; L40888; AAC41598.1; -; Genomic_DNA.
  FT NON TER 1
  SQ SEQUENCE 280 AA; 32944 MW; 467085E23BABBC27 CRC64;

Query Match 91.7%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.2e-37;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 RKADTKKLERKKEHGDILAEADLYGRLEIPALPELPSENERGYIIPHOSSLPODNRGNSRD 70
```

Db 1 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHQSLPDNRGNSRD 60

Qy 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 117

Db 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 10

Q25851_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25851;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxId=5833;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."
RT Mol. Biochem. Parasitol. 71:291-294(1995).
RL EMBL; L40893; AAC41603.1; -; Genomic_DNA.
RT NON_TER 1
SQ SEQUENCE 280 AA; 32927 MW; 48706EPDCCF40ADD3 CRC64;

Query Match 91.7%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.2e-37;

Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHQSLPDNRGNSRD 70

Db 1 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHQSLPDNRGNSRD 60

Qy 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 117

Db 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 11

Q25848_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25848;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxId=5833;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."
RT Mol. Biochem. Parasitol. 71:291-294(1995).
RL EMBL; L40889; AAC41599.1; -; Genomic_DNA.
RT NON_TER 1
SQ SEQUENCE 280 AA; 32939 MW; E77080385FPA0D61 CRC64;

Query Match 91.2%; Score 550; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 2.1e-37;

Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHQSLPDNRGNSRD 70

Db 1 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHQSLPDNRGNSRD 60

Qy 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 117

Db 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 12

Q25852_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25852;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxId=5833;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."
RT Mol. Biochem. Parasitol. 71:291-294(1995).
RL EMBL; L40834; AAB59232.1; -; Genomic_DNA.
RT NON_TER 1
SQ SEQUENCE 280 AA; 32927 MW; E6648F85FBBA245 CRC64;

Query Match 91.2%; Score 550; DB 2; Length 280;
Best Local Similarity 98.1%; Pred. No. 2.1e-37;

Matches 105; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHQSLPDNRGNSRD 70

Db 1 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHQSLPDNRGNSRD 60

Qy 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 117

Db 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 13

Q25845_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25845;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxId=5833;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."
RT Mol. Biochem. Parasitol. 71:291-294(1995).
RL EMBL; L40886; AAB59232.1; -; Genomic_DNA.
RT NON_TER 1
SQ SEQUENCE 280 AA; 32908 MW; FDEE929F34DCA7E9 CRC64;

Query Match 91.0%; Score 549; DB 2; Length 280;
 Best Local Similarity 98.1%; Pred. No. 2.5e-37;
 Matches 105; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

11 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 70
 1 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 60

QY 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
 DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107

RESULT 14
 Q25853_PLAFA
 ID Q25853_PLAFA PRELIMINARY; PRT; 280 AA.
 AC Q25853

DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Liver stage-specific antigen 1 (Fragment).

GN Name=LSA-1;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_Taxid=5833;
 RN 1]

RP NUCLEOTIDE SEQUENCE.
 RA MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
 RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
 Hawley W.A., Collins W.E., Lal A.A.,

RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."
 RT Mol. Biochem. Parasitol. 71:291-294(1995).

DR EMBL: I40835; AAB59234.1; -, Genomic_DNA.
 FT NON_TER 1 1
 SQ SEQUENCE 280 AA; 32938 MW; 46751C45F4DCBD33 CRC64;

Query Match 91.0%; Score 549; DB 2; Length 280;
 Best Local Similarity 98.1%; Pred. No. 2.5e-37;
 Matches 105; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 70
 DB 1 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 60

QY 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
 DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107

RESULT 15

Q25854_PLAFA
 ID Q25854_PLAFA PRELIMINARY; PRT; 280 AA.
 AC Q25854

DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Liver stage-specific antigen 1 (Fragment).

GN Name=LSA-1;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_Taxid=5833;
 RN 1]

RP NUCLEOTIDE SEQUENCE.
 RA MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
 RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
 Hawley W.A., Collins W.E., Lal A.A.,

RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."
 RT Mol. Biochem. Parasitol. 71:291-294(1995).

DR EMBL: I40836; AAC1604.1; -, Genomic_DNA.

FT NON_TER 1 1
 SQ SEQUENCE 280 AA; 32966 MW; 61851C5ADE1A211 CRC64;

Query Match 91.0%; Score 549; DB 2; Length 280;
 Best Local Similarity 98.1%; Pred. No. 2.5e-37;
 Matches 105; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 70
 DB 1 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 60

QY 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
 DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107

Search completed: December 7, 2005, 07:20:44
 Job time : 39.0185 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 06:57:58 ; Search time 37.8291 Seconds
(without alignment)
1358.936 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603 1 LDEEQRDLEQRKADTKKLE.....GLHEKKKGSIRPEKKEDKS 117

Sequence: 1 LDEEQRDLEQRKADTKKLE.....GLHEKKKGSIRPEKKEDKS 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21.*
1: geneseqp19808.*
2: geneseqp19908.*
3: geneseqp20008.*
4: geneseqp20018.*
5: geneseqp20028.*
6: geneseqp20038.*
7: geneseqp20038.*
8: geneseqp20048.*
9: geneseqp20058.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	603	100.0	117	2	AAR26937 P.falcipa
2	603	100.0	316	2	AAR26941 P.falcipa
3	603	100.0	493	2	AAR26944 P.falcipa
4	602	99.8	456	8	ADO21942 LSA-NRC(H
5	602	99.8	457	8	ADO21920 LSA-NRC(H
6	172	28.5	34	8	ADO21926 Malaria p
7	103	17.1	20	8	ADO21925 Malaria p
8	94	15.6	452	9	Ad240617 PICTL/HTL
9	94	15.6	456	7	Ad240617 Multi-epi
10	94	15.6	456	8	Ado24115 Epigene c
11	90.5	15.0	1085	2	AAR95607 RRP3 telo
12	90.5	15.0	1085	7	ADK62464 Disease t
13	89.5	14.8	231	4	AAG98701 Human cel
14	89.5	14.8	232	4	AAG98700 Human cel
15	89.5	14.8	238	4	AAG98699 Human cel
16	89.5	14.8	257	4	AAG98698 Human cel
17	89.5	14.8	272	4	AAG98697 Human cel
18	87.5	14.5	271	5	ABB04736 Human cyl
19	87.5	14.5	282	3	AAY58635 Protein r
20	86	14.3	550	8	ADN23342 Bacterial
21	86	14.3	447	4	AAG82903 S. epider
22	85.5	14.2	450	4	AAG82906 S. epider
23	85.5	14.2	684	8	Ad116245 Human nuc
24	85.5	14.2	765	6	ABM35013 Cancer ba

25	85.5	14.2	765	6	ADB20866 MKP1 baee
26	85.5	14.2	765	6	ABB82815 Human top
27	85.5	14.2	765	7	ADB87955 Human UGT
28	85.5	14.2	765	7	ADB96938 Human MDR
29	85.5	14.2	765	7	ADB92129 Human MDR
30	85.5	14.2	765	7	ADJ69343 Human hea
31	85.5	14.2	765	8	ADJ66581 DNA topol
32	85.5	14.2	765	8	ADL82893 Human PRO
33	85.5	14.2	765	8	ADP56068 Human PRO
34	85	14.1	15	2	AAW85415 Helper T-
35	85	14.1	15	2	AAW85172 Helper T-
36	85	14.1	15	7	Adw36112 HLA bindi
37	85	14.1	15	7	ADW35943 HLA bindi
38	84.5	14.0	150	5	ABB08082 Human top
39	84.5	14.0	242	4	ABG02495 Novel hum
40	84.5	14.0	300	5	ABB08081 Human top
41	84.5	14.0	323	8	ADN04692 Antisport
42	84.5	14.0	585	8	ADJ75618 Marker ge
43	84.5	14.0	585	1	ADX06678 Cyclin-de
44	84.5	14.0	765	1	AAp92275 Human top
45	84.5	14.0	765	2	AAR75915 Variant h

ALIGNMENTS

RESULT 1
AAR26937
ID AAR26937 standard; peptide; 117 AA.
XX
AC AAR26937;
XX
XX 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX
DE P.falciparum liver-stage specific antigen non-repeat region.
XX
XX Malaria; LSA-NR; hepatocyte; sporozoite; plasmodium DG 536; T-cell epitope;
XX paludism.
XX
XX Plasmodium falciparum.
OS
XX WO9213884-A1.
PN
XX 20-AUG-1992.
PD
XX 05-FEB-1992; 92WO-FR000104.
PF
XX 05-FEB-1991; 91FR-00001286.
PR
XX (INSP) INST PASTEUR.
XX
XX Guerinmarchand C, Drullhe P;
PI WPI; 1992-299985/36.
PT Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.
XX
XX Claim 4; Page 55; 81pp; French.
XX
XX A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
XX gIII was used to transform E.coli. The expression library was screened
XX with human antisera against antigens of all stages of P. falciparum. The
XX library was rescreened with antibodies affinity- purified on a clone
XX which was able to recognise antibodies specific to the hepatic phase.
XX About 40 clones were detected which produced a characteristic LSA
XX epitope. The clone with the largest insert encoded LSA-R-NR containing a
XX 12-repeat region (LSA-R) followed by a non-repeat region (LSA-NR). The
XX sequence given here is a preferred polypeptide of the invention carrying
XX a T cell epitope typical of liver-stage P.falciparum. See AAR26919-
XX R26935, AAR26937 and AAR26939. (Updated on 25-MAR-2003 to correct PN
XX field.)

```
XX Sequence 117 AA;
SQ
Query Match 100.0%; Score 603; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 8.3e-57;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEQORDLEQKADTKKULERKKEHGDIADLYGRLEIPALIELPSENERGYIIPHOSSL 60
DB 1 LOEQORDLEQKADTKKULERKKEHGDIADLYGRLEIPALIELPSENERGYIIPHOSSL 60

OY 61 PODNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
DB 61 PODNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117

RESULT 2
AAR26941 standard; protein; 316 AA.
XX
AC AAR26941;
XX
DT 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX
DE P.falciparum LSA-R-NR protein.
XX
KW Malaria; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;
KW paludism; liver stage-specific antigen.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT 1..209
FT /label= repeat_region
FT /note= "contains 12 x 17mer repeats"
FT 210..316
FT Region /label= non-repeat_region
XX
FN WO9213884-A1.
XX
PD 20-AUG-1992.
XX
PF 05-FEB-1992; 92WO-FR000104.
XX
PR 05-FEB-1991; 91FR-00001286.
XX
PA (INSP ) INST PASTEUR.
XX
PI Guerimarchand C, Druilhe P;
XX
DR WPI, 1992-299985/36.
DR N-PSDB; AAQ28115.
XX
PT Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.
XX
PS Disclosure; Fig 1; 81pp; French.
XX
CC A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
CC gtl1 was used to transform E.coli. The expression library was screened
CC with human antisera against antigens of all stages of P.falciparum. The
CC library was rescreened with antibodies affinity-purified on a clone
CC which was able to recognise antibodies specific to the hepatic phase.
CC About 40 clones were detected which produced a characteristic LSA
CC epitope. The clone with the largest insert (950 bases) encoded LSA-R-NR
CC containing a 12-repeat region followed by a non-repeat region. Preferred
CC antigenic polypeptides of the invention are derived from the amino acid
CC sequence of LSA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 316 AA;
Query Match 100.0%; Score 603; DB 2; Length 316;
```

```
Best Local Similarity 100.0%; Pred. No. 3.1e-56;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEQORDLEQKADTKKULERKKEHGDIADLYGRLEIPALIELPSENERGYIIPHOSSL 60
DB 200 LOEQORDLEQKADTKKULERKKEHGDIADLYGRLEIPALIELPSENERGYIIPHOSSL 259

OY 61 PODNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
DB 260 PODNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 316

RESULT 3
AAR26944 standard; protein; 493 AA.
XX
AC AAR26944;
XX
DT 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX
DE P.falciparum LSA gene C-terminal region.
XX
KW Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
KW paludism; liver stage-specific antigen.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT 13..213
FT /label= repetitive_region
FT 214..493
FT Region /label= non-repetitive_region
XX
FN WO9213884-A1.
XX
PD 20-AUG-1992.
XX
PF 05-FEB-1992; 92WO-FR000104.
XX
PR 05-FEB-1991; 91FR-00001286.
XX
PA (INSP ) INST PASTEUR.
XX
PI Guerimarchand C, Druilhe P;
XX
DR WPI, 1992-299985/36.
DR N-PSDB; AAQ28119.
XX
PT Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.
XX
PS Claim 2; Fig 8-10; 81pp; French.
XX
CC The 3' part of the P.falciparum liver-stage specific antigen (LSA) gene
CC codes for a polypeptide sequence which carries a T cell epitope
CC characteristic of a protein produced in hepatocytes infected with
CC P.falciparum. The polypeptide can be used in the preparation of vaccines
CC against malaria. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 493 AA;
Query Match 100.0%; Score 603; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 5.5e-56;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEQORDLEQKADTKKULERKKEHGDIADLYGRLEIPALIELPSENERGYIIPHOSSL 60
DB 204 LOEQORDLEQKADTKKULERKKEHGDIADLYGRLEIPALIELPSENERGYIIPHOSSL 263

OY 61 PODNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
DB 264 PODNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 320
```

```
RESULT 4
ADO21942
ID ADO21942 standard; protein; 456 AA.
XX
AC ADO21942;
XX
DT 12-AUG-2004 (first entry)
XX
DE LSA-NRC(H) construct protein derived from Malaria parasite LSA-1.
XX
XX LSA-NRC, liver stage antigen; LSA-1; T-cell; B-cell epitope;
XX parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
XX malaria parasite P. falciparum; LSA-NRC(H) construct.
XX
OS Plasmodium falciparum.
OS Synthetic.
XX
PN WO2004044167-A2.
XX
PD 27-MAY-2004.
XX
PF 12-NOV-2003; 2003WO-US036011.
XX
PR 12-NOV-2002; 2002US-0425719P.
XX
PA (REED-) REED ARMY INST RES WALTER.
XX
PI Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Barbosa A;
XX
DR WPI; 2004-420309/39.
DR N-PSDB; ADO21941.
XX
PT Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
PS Claim 4; SEQ ID NO 26; 90pp; English.
XX
CC The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct protein of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to a C-terminal His6 tag.
XX
SQ Sequence 456 AA;
XX
Query Match 99.8%; Score 602; DB 8; Length 456;
Best Local Similarity 99.1%; Pred. No. 6,4e-56;
Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LQEQQDLEQRKADTKNLERKKEHGDIIAEDLYGRLEIPALPSPENRGYYIPIHQSSL 60
DB 155 LQEQQDLEQRKADTKNLERKKEHGDVLAEDLYGRLEIPALPSPENRGYYIPIHQSSL 214
QY 61 PQQNRGNSRDSKEISIIETKTNRESITTNVGRDIIHGHLSEKKDGSIKPEQKEDKS 117
DB 215 PQQNRGNSRDSKEISIIETKTNRESITTNVGRDIIHGHLSEKKDGSIKPEQKEDKS 271
XX
RESULT 5
ADO21920
ID ADO21920 standard; protein; 457 AA.
XX
AC ADO21920;
XX
```

```
DT 12-AUG-2004 (first entry)
XX
DE LSA-NRC(H)Mut construct protein derived from Malaria parasite LSA-1.
XX
XX LSA-NRC, liver stage antigen; LSA-1; T-cell; B-cell epitope;
XX parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
XX malaria parasite P. falciparum; LSA-NRC(H)Mut construct; mutant; mtein.
XX
OS Plasmodium falciparum.
OS Synthetic.
XX
PN WO2004044167-A2.
XX
PD 27-MAY-2004.
XX
PF 12-NOV-2003; 2003WO-US036011.
XX
PR 12-NOV-2002; 2002US-0425719P.
XX
PA (REED-) REED ARMY INST RES WALTER.
XX
PI Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Barbosa A;
XX
DR WPI; 2004-420309/39.
DR N-PSDB; ADO21919.
XX
PT Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
PS Claim 5; SEQ ID NO 4; 90pp; English.
XX
CC The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct protein of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to an amino acid insertion
CC within the TS P. falciparum epitope and a C-terminal His6 tag.
XX
SQ Sequence 457 AA;
XX
Query Match 99.8%; Score 602; DB 8; Length 457;
Best Local Similarity 99.1%; Pred. No. 6,4e-56;
Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LQEQQDLEQRKADTKNLERKKEHGDIIAEDLYGRLEIPALPSPENRGYYIPIHQSSL 60
DB 155 LQEQQDLEQRKADTKNLERKKEHGDVLAEDLYGRLEIPALPSPENRGYYIPIHQSSL 214
QY 61 PQQNRGNSRDSKEISIIETKTNRESITTNVGRDIIHGHLSEKKDGSIKPEQKEDKS 117
DB 215 PQQNRGNSRDSKEISIIETKTNRESITTNVGRDIIHGHLSEKKDGSIKPEQKEDKS 271
XX
RESULT 6
ADO21926
ID ADO21926 standard; peptide; 34 AA.
XX
AC ADO21926;
XX
DT 12-AUG-2004 (first entry)
XX
```

DE Malaria parasite P. falciparum LSA-1 LSA-ter epitope peptide.
XX
XX LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
KM parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
XX Malaria parasite P. falciparum; epitope; LSA-ter.
XX Plasmodium falciparum.
OS
PN WO200404167-A2.
PD
PD 27-MAY-2004.
PF 12-NOV-2003; 2003WO-US036011.
PR 12-NOV-2002; 2002US-0425719P.
XX
XX (REED-) REED ARMY INST RES WALTER.
PA
PI Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Barboea A;
XX
XX WPI; 2004-420309/39.
DR
XX
XX Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
XX Claim 2; SEQ ID NO 10; 90pp; English.
PS
XX
XX The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the malaria parasite P. falciparum LSA-1 LSA-ter epitope
CC peptide of the invention.
XX
XX Sequence 34 AA;
SQ
Query Match 28.5%; Score 172; DB 8; Length 34;
Best Local Similarity 100.0%; Pred. No. 4,4e-11;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 67 NSRDSKEISITKTNRESITTNVEGRDIIHGHL 100
DB 1 NSRDSKEISITKTNRESITTNVEGRDIIHGHL 34
RESULT 7
ADO21925
ID ADO21925 standard; peptide; 20 AA.
XX
XX ADO21925;
AC
DT 12-AUG-2004 (first entry)
XX
XX Malaria parasite P. falciparum liver stage antigen-1 NR epitope peptide.
DE
XX LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
KM parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
KW malaria parasite P. falciparum; epitope; NR.
XX Plasmodium falciparum.
OS
PN WO200404167-A2.
PD
PD 27-MAY-2004.
PF 12-NOV-2003; 2003WO-US036011.
XX
XX 12-NOV-2002; 2002US-0425719P.
PR
XX

PA (REED-) REED ARMY INST RES WALTER.
XX
XX Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Barboea A;
XX
XX WPI; 2004-420309/39.
DR
XX
XX Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
XX Claim 2; SEQ ID NO 9; 90pp; English.
PS
XX
XX The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the malaria parasite P. falciparum LSA-1 NR epitope peptide of
CC the invention.
XX
XX Sequence 20 AA;
SQ
Query Match 17.1%; Score 103; DB 8; Length 20;
Best Local Similarity 95.0%; Pred. No. 0.00059;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 15 TKKNLERKKEHGDVLABDLY 34
DB 1 TKKNLERKKEHGDVLABDLY 20
RESULT 8
ADZ40617
ID ADZ40617 standard; protein; 452 AA.
XX
XX ADZ40617;
AC
DT 30-JUN-2005 (first entry)
XX
XX
XX pFCTT/HTL/N multi-epitope construct protein.
DE
XX multi-epitope; immune stimulation; HIV infection; anti-HIV;
KW immunostimulant; vaccine; immunogenicity.
XX
XX Synthetic.
OS
PN WO2005033265-A2.
PD
PD 14-APR-2005.
PF 26-APR-2004; 2004WO-US012732.
XX
XX 25-APR-2003; 2003US-0465229P.
PR
XX (EPIM-) EPIMUNE INC.
PA
PI Sette A, Chesnut RW, Newman MJ, Livingston BD;
PI WPI; 2005-285416/29.
DR
DR N-PSDB; ADZ40618.
XX
XX New multi-epitope polynucleotide, useful for designing vaccines that
PT provide increased immunogenicity, and for inducing immune responses
PT against HIV in an individual.
XX
XX Disclosure; Fig 181; 261pp; English.
PS
XX
XX The invention relates to a novel multi-epitope polynucleotide. The
CC invention further comprises: a composition comprising the polynucleotide;
CC a cell comprising the polynucleotide; a method for inducing an immune
CC response against human immunodeficiency virus (HIV) in an individual,

CC comprising administering the polynucleotide, composition or cell to the
CC individual; and a method for making the polynucleotide, composition or
CC cell. The multi-epitope polynucleotide has anti-HIV and immunostimulant
CC activities. The polynucleotide is useful for designing vaccines that
CC provide increased immunogenicity, and for inducing an immune response
CC against HIV. This sequence represents an pCTL/HTL/N multi-epitope
CC construct protein of the invention.

XX Sequence 452 AA;
SQ

Query Match 15.6%; Score 94; DB 9; Length 452;
Best Local Similarity 85.7%; Pred. No. 0.33;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 40 PAIELPSENERGYIIPHOSSL 60
DB 119 PGPGLPSENERGYIIPHOSSL 139

RESULT 9
ADA49437
ID ADA49437 standard; protein; 456 AA.
XX
AC ADA49437;
XX
DT 20-NOV-2003 (first entry)
XX
DE Multi-epitope construct #21.
XX
KM multi-epitope; immunogenic; epitope; major histocompatibility complex;
XX MHC class I; MHC class II; junctional epitope.
XX
OS Synthetic.
XX Plasmidium falciparum.
XX
PN US2002119127-A1.
XX
PD 29-AUG-2002.
XX
PF 27-JUN-2001; 2001US-00894018.
XX
PR 28-DEC-1999; 99US-0173390P.
XX 28-DEC-2000; 2000MO-US035568.
XX 16-APR-2001; 2001US-0284221P.
XX
PA (SETT/) SETTE A.
XX (CHES/) CHESNUT R.
XX (LIIV/) LIVINGSTON B D.
XX (BAKE/) BAKER D M.
XX (NEWM/) NEWMAN M J.
XX (BROW/) BROWN D H.
XX
PI Sette A, Chesnut R, Livingston BD, Baker DM, Newman MJ, Brown DH;
XX WPI, 2003-615704/58.
XX N-PSDB; ADA49438.
XX
PT Designing multi-epitope construct having major histocompatibility complex
XX class I and II epitope nucleic acids, by selecting mixture of amino acid
XX insertions at junctions of construct to minimize junctional epitopes.
XX
PS Disclosure; Fig 18f; 78pp; English.
XX
XX The invention relates to a method of designing multi-epitope constructs
CC comprising major histocompatibility complex (MHC) class I and II (CTL)
CC epitope nucleic acids (CEN), involves sorting CEN, introducing flanking
CC amino acid residue selected from specified amino acid residues given in
CC specification at C+1 position of CEN, introducing amino acid spacer
CC residues between two CEN, and selecting the constructs having less
CC junctional epitopes. The method is useful for designing a multi-epitope
CC construct having multiple nucleic acid. The method avoids or
CC minimizes the occurrence of junctional epitopes and maximizes the
CC immunogenicity and/or antigenicity of multi-epitope vaccines. The present

CC sequence represents the amino acid sequence of a multi-epitope construct.
XX
SQ Sequence 456 AA;
SQ

Query Match 15.6%; Score 94; DB 7; Length 456;
Best Local Similarity 85.7%; Pred. No. 0.34;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 40 PAIELPSENERGYIIPHOSSL 60
DB 119 PGPGLPSENERGYIIPHOSSL 139

RESULT 10
AD024115
ID AD024115 standard; protein; 456 AA.
XX
AC AD024115;
XX
DT 01-JUL-2004 (first entry)
XX
DE Epigene construct pCTL/HTL(N) protein.
XX
KM hepatotropic; virucide; anti-inflammatory; anti-HIV; cytostatic; epitope;
XX cytotoxic T lymphocyte; CTL; helper T lymphocyte; HTL; immune response;
XX hepatitis B virus; HBV; HIV; cancer; lymphoma.
XX
OS Unidentified.
XX
PN WO2004031210-A2.
XX
PD 15-APR-2004.
XX
PF 03-OCT-2003; 2003WO-US031303.
XX
PR 03-OCT-2002; 2002US-0415463P.
XX 22-OCT-2002; 2002US-0419973P.
XX
PA (EPIM-) EPIMUNE INC.
XX (GENV-) GENENCOR INT INC.
XX
PI Sette A, Chesnut R, Newman MJ, Livingston BD, Babe LM, Chen Y;
XX Deyoung LM, Huang MTF, Power SD;
XX WPI, 2004-330143/30.
XX N-PSDB; AD024116.
XX
PT New multi-epitope polynucleotides encoding cytotoxic T lymphocyte and/or
XX helper T lymphocyte epitopes, useful for inducing or stimulating an
XX antiviral or anticancer immune response, especially against hepatitis B
XX virus.
XX
PS Disclosure; SEQ ID NO 265; 401pp; English.
XX
XX The invention relates to a multi-epitope polynucleotide (I) encoding
CC cytotoxic T lymphocyte (CTL) and/or helper T lymphocyte (HTL) epitopes.
CC The composition and methods are useful for inducing or stimulating an
CC immune response against hepatitis B virus (HBV), other viruses (e.g.
CC HIV), or against cancers such as prostate cancer or lymphoma. This
CC sequence corresponds to an example of an epigene construct encoded
CC protein used in the invention.
XX
SQ Sequence 456 AA;
SQ

Query Match 15.6%; Score 94; DB 8; Length 456;
Best Local Similarity 85.7%; Pred. No. 0.34;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 40 PAIELPSENERGYIIPHOSSL 60
DB 119 PGPGLPSENERGYIIPHOSSL 139

RESULT 11

AAR95607

ID AAR95607 standard; protein; 1085 AA.

XX AAR95607;

XX 12-OCT-1996 (first entry)

XX RRP3 telomerase-associated protein.

XX Yeast; RRP3 protein; telomerase-associated protein; STR7;

XX suppressor of telomeric repression-7; telomerase; ribonucleoprotein;

XX telomere; tumour; pathogen; sperm; ovum; reporter gene; drug screening;

XX antibody; immunoassay; antitumour; antiseptic; contraceptive;

XX interility; diagnostic; gene therapy.

XX Saccharomyces cerevisiae.

XX WO9612811-A2.

XX 02-MAY-1996.

XX 20-OCT-1995; 95WO-US013801.

XX 20-OCT-1994; 94US-00326781.

XX 28-APR-1995; 95US-00431080.

XX (ARCH-) ARCH DEV CORP.

XX Gottschling DE, Singer MS;

XX WPI; 1996-239169/24.

XX N-PSDB; AAT27052.

XX Novel telomerase associated polypeptide(s) and related nucleic acid -

XX useful for detecting e.g. tumour cells or pathogens.

XX Example 10; Page 294-302; 349pp; English.

XX RRP3 telomerase-associated protein (containing fragment STR7 (AAR95606)

XX is found in conjunction with Saccharomyces cerevisiae telomerase, a

XX ribonucleoprotein required for telomere replication. Other proteins

XX associated with telomerase are given in AAR95601-05 (other STR proteins).

XX These proteins combine with telomerase to repress telomere silencing of

XX gene expression. Oligonucleotides from the encoding sequence may be used

XX to detect non-ciliate telomerase-associated genes, e.g. in tumour, may be

XX pathogen, sperm or ovum cells. New telomerase-associated sequences may be

XX detected by a reporter gene expression system linked to an expression.

XX representing telomere sequence, and binding compounds, e.g. antibodies, may

XX be detected by complex formation with telomerase components. The products

XX may be used as antitumour, antiseptic or contraceptive agents, in

XX interility diagnosis, or in gene therapy

XX Sequence 1085 AA;

SQ

Query Match 15.0%; Score 90.5; DB 2; Length 1085;

Best Local Similarity 20.8%; Pred. No. 2.5;

Matches 31; Conservative 21; Mismatches 54; Indels 43; Gaps 4;

DB 3 EQRDLQRKADTKNLERKKE-----HGD 27

DB 865 DKMKELQRKTEKKRQLKKKKLLKIRKQROKAIKEQETMNLQLGINGHEITGNNSHD 924

DB 28 I-LAEDLYGRLEIPALPESE-NEGGYIIPHOSSIPODNRGNSRDSKEI-----SIEEK 79

DB 925 INTGDTFTNTENTPMNELPSHAPEDASLIPNNSDLAVSNTRKSTKSVGDEIHEILGK 984

DB 80 TNRESITTNVEGRDIIHKHLEKKDGS 108

DB 985 DENDLLSVGDIINGYDAQEGHVIETDADI 1013

RESULT 12

ADK62464

ID ADK62464 standard; protein; 1085 AA.

XX ADK62464;

XX 06-MAY-2004 (first entry)

XX Disease treating protein complex-derived protein #358.

XX protein complex; drug target; diagnosis.

XX Unidentified.

XX EPI338608-A2.

XX 27-AUG-2003.

XX 20-DEC-2002; 2002EP-00102902.

XX 20-DEC-2001; 2001EP-00130253.

XX (CELL-) CELZOME AG.

XX Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;

XX Marzloch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;

XX Michon A, Leutwein C, Rick J;

XX WPI; 2003-638460/61.

XX N-PSDB; ADK62465.

XX New proteins and protein complexes from eukaryotes, useful as targets in

XX drug screening, or in diagnosing or screening for the presence of a

XX disease or disorder, or a predisposition for developing a disease or

XX disorder in a subject.

XX Disclosure; SEQ ID NO 715; 13pp; English.

XX The invention relates to novel protein complexes comprising a first and a

XX second protein, or its derivative, fragment, homologue or variant. The

XX proteins are selected from given protein complexes, which are not defined

XX in the specification. The variants are encoded by nucleic acids that

XX hybridize to the nucleic acids encoding the proteins under low stringency

XX conditions. The protein complexes are useful as targets for an active

XX agent of a pharmaceutical. These protein complexes are particularly

XX useful as drugs targets for the treatment or preventing of a disease or

XX disorder. The complexes and methods above are useful in diagnosing or

XX screening for the presence of a disease or disorder or a predisposition

XX for developing a disease or disorder in a subject. These are also useful

XX in screening for a drug for treatment or prevention of a disease or

XX disorder. The molecule that modulates the amount, activity or protein

XX components of the complex is useful for the manufacture of a medicament

XX for the treatment or prevention of a disease or disorder. This sequence

XX corresponds to a protein of the invention. (Note: the sequence data for

XX this patent did not form part of the printed specification but was

XX obtained from the EPO in electronic format).

SQ

Query Match 15.0%; Score 90.5; DB 7; Length 1085;

Best Local Similarity 20.8%; Pred. No. 2.5;

Matches 31; Conservative 21; Mismatches 54; Indels 43; Gaps 4;

DB 3 EQRDLQRKADTKNLERKKE-----HGD 27

DB 865 DKMKELQRKTEKKRQLKKKKLLKIRKQROKAIKEQETMNLQLGINGHEITGNNSHD 924

DB 28 I-LAEDLYGRLEIPALPESE-NEGGYIIPHOSSIPODNRGNSRDSKEI-----SIEEK 79

DB 925 INTGDTFTNTENTPMNELPSHAPEDASLIPNNSDLAVSNTRKSTKSVGDEIHEILGK 984

DB 80 TNRESITTNVEGRDIIHKHLEKKDGS 108

DB 985 DENDLLSVGDIINGYDAQEGHVIETDADI 1013

RESULT 13
AAG98701
ID AAG98701 standard; protein; 231 AA.
XX
AC AAG98701;
XX
DT 21-SEP-2001 (first entry)
XX
DE Human cell death protective cDNA clone CNI-00717 ORF5 protein, SEQ:194.
XX
KW Cell death protective; apoptosis; necrosis; human; drug screening;
KW cell death-associated disorder; central nervous system disorder;
KW psychiatric disorder; neurological disorder; ischaemia-related disorder;
KW stroke; cerebral infarction; ischaemic encephalopathy;
KW neurodegenerative disorder; Alzheimer's disease; Huntington's disease;
KW Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;
KW vascular disease; ophthalmological disorder; diabetic retinopathy;
KW macular degeneration; hypertension; myocardial infarction;
KW atherosclerosis; respiratory disorder; asthma; transgenic animal;
KW chronic obstructive pulmonary disease; neoplastic condition; cancer;
KW benign tumour; anaemia; gastrointestinal disorder; gastritis;
KW ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;
KW glomerulonephritis; cystitis; endometriosis; endocrine disorder;
KW Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;
KW urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS.
XX
OS Homo sapiens.
XX
PN WO200145638-A2.
XX
PD 28-JUN-2001.
XX
PF 11-DEC-2000; 2000WO-US033547.
XX
PR 14-DEC-1999; 99US-00461697.
XX
PA (COGE-) COGENT NEUROSCIENCE INC.
XX
PI Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;
XX
DR WPI: 2001-390297/41.
XX
N-PSDB; AAH84226, AAH84231.
XX
PT Novel protective sequence polynucleotides and polypeptides, used to
PT identify modulators of their expression and activity, which are used in
PT to treat central nervous system conditions, diseases and disorders.
XX
PS Claim 1; Fig 9E; 325pp; English.
XX
XX Sequences AAH84132-AAH84370 represent human nucleic acid sequences which
XX protect against cell death (i.e., apoptosis or necrosis). Sequences
XX AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,
XX AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,
XX while the remaining nucleic acid sequences within the range given above
XX represent the open reading frames (ORFs) of these cDNA clones. Sequences
XX AAG98610-AAG98829 represent the polypeptides encoded by the cell death
XX protective ORFs. The cell death protective cDNA clones are able to
XX prevent, delay or reverse progression through the apoptotic or necrotic
XX pathways when injected into a cell predisposed to or undergoing cell
XX death. The cell death protective nucleic acids and polypeptides can be
XX used in the diagnosis and treatment of disorders associated with cell
XX death, and to screen for compounds which modulate their activity or
XX expression. Such modulators, preferably a small organic molecule, an
XX antibody, a ribozyme, or an antisense molecule, can also be used to treat
XX cell death-related diseases. Such diseases include those associated with
XX the central nervous system including psychiatric or neurological
XX disorders, especially ischaemia-related conditions such as strokes, and
XX also includes neurodegenerative disorders such as Alzheimer's disease,
XX Huntington's disease, or Parkinson's disease. The modulators may also be
XX used to treat infections such as meningitis, malaria, or trypanosomiasis;
XX vascular diseases such as ischaemic encephalopathy or cerebral infarction

CC : eye conditions such as diabetic retinopathy or macular degeneration;
CC hypertension; myocardial infarction; atherosclerosis; respiratory
CC conditions such as asthma or chronic obstructive pulmonary disease;
CC neoplastic conditions such as cancers or benign tumours; blood cell
CC conditions such as anaemia; gastrointestinal conditions such as gastritis
CC or ulcerative colitis; liver conditions such as biliary cirrhosis; kidney
CC disorders such as glomerulonephritis; cystitis; endometriosis; endocrine
CC disorders such as Grave's disease or Hashimoto's thyroiditis; skin
CC conditions such as dermatitis or urticaria; or immune system disorders
CC such as acquired immunodeficiency syndrome (AIDS). The nucleic acids may
CC additionally be used to generate animal models of cell death-associated
CC disorders. The present sequence represents a cell death protective
CC polypeptide
XX
SQ Sequence 231 AA;
XX
Query Match 14.8%; Score 89.5; DB 4; Length 231;
Best Local Similarity 29.3%; Pred. No. 0.42;
Matches 34; Conservative 15; Mismatches 60; Indels 7; Gaps 3;
QY 3 EQQDLEQRKADTKNLRKKKEHGDILADLYGRLEIFALRPSNENGYIYPHQ--SL 60
DB 71 EVKNEEEDQKEDBEDQNEKEGAG--KEDXDEKEBEDQKEDKXNGNEGDAKEDGKK 127
QY 61 PQDNRGNRSDSKSEISIEKTRRESITTNV--EGRRDIHKHGLEEKDGSIRPEQKE 114
DB 128 GEDGNGNBDGKEKEDKEBEDRKETGVGENEDGKEKGDKKGKVYKEDKE 183
RESULT 14
AAG98700
ID AAG98700 standard; protein; 232 AA.
XX
AC AAG98700;
XX
DT 21-SEP-2001 (first entry)
XX
DE Human cell death protective cDNA clone CNI-00717 ORF4 protein, SEQ:192.
XX
KW Cell death protective; apoptosis; necrosis; human; drug screening;
KW cell death-associated disorder; central nervous system disorder;
KW psychiatric disorder; neurological disorder; ischaemia-related disorder;
KW stroke; cerebral infarction; ischaemic encephalopathy;
KW neurodegenerative disorder; Alzheimer's disease; Huntington's disease;
KW Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;
KW vascular disease; ophthalmological disorder; diabetic retinopathy;
KW macular degeneration; hypertension; myocardial infarction;
KW atherosclerosis; respiratory disorder; asthma; transgenic animal;
KW chronic obstructive pulmonary disease; neoplastic condition; cancer;
KW benign tumour; anaemia; gastrointestinal disorder; gastritis;
KW ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;
KW glomerulonephritis; cystitis; endometriosis; endocrine disorder;
KW Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;
KW urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS.
XX
OS Homo sapiens.
XX
PN WO200145638-A2.
XX
PD 28-JUN-2001.
XX
PF 11-DEC-2000; 2000WO-US033547.
XX
PR 14-DEC-1999; 99US-00461697.
XX
PA (COGE-) COGENT NEUROSCIENCE INC.
XX
PI Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;
XX
DR WPI: 2001-390297/41.
XX
N-PSDB; AAH84226, AAH84230.
XX
PT Novel protective sequence polynucleotides and polypeptides, used to

PT identify modulators of their expression and activity, which are used in
PT to treat central nervous system conditions, diseases and disorders.
PS Claim 1, Fig 9D, 325pp; English.
XX Sequences AAH84132-AAH84370 represent human nucleic acid sequences which
CC protect against cell death (i.e., apoptosis or necrosis). Sequences
CC AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,
CC AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,
CC while the remaining nucleic acid sequences within the range given above
CC represent the open reading frames (ORFs) of these cDNA clones. Sequences
CC AAH98610-AAH98829 represent the polypeptides encoded by the cell death
CC protective ORFs. The cell death protective cDNA clones are able to
CC prevent, delay or reverse progression through the apoptotic or necrotic
CC pathways when injected into a cell predisposed to or undergoing cell
CC death. The cell death protective nucleic acids and polypeptides can be
CC used in the diagnosis and treatment of disorders associated with cell
CC death, and to screen for compounds which modulate their activity or
CC expression. Such modulators, preferably a small organic molecule, an
CC antibody, a ribozyme, or an antisense molecule, can also be used to treat
CC cell death-related diseases. Such diseases include those associated with
CC the central nervous system including psychiatric or neurological
CC disorders, especially ischemia-related conditions such as strokes, and
CC also includes neurodegenerative disorders such as Alzheimer's disease,
CC Huntington's disease, or Parkinson's disease. The modulators may also be
CC used to treat infections such as ischemic encephalopathy or cerebral infarction;
CC vascular diseases such as diabetic retinopathy or macular degeneration;
CC hypertension; myocardial infarction; atherosclerosis; respiratory
CC conditions such as asthma or chronic obstructive pulmonary disease;
CC neoplastic conditions such as cancers or benign tumours; blood cell
CC conditions such as anaemia; gastrointestinal conditions such as gastritis
CC or ulcerative colitis; liver conditions such as biliary cirrhosis; kidney
CC disorders such as glomerulonephritis; cystitis; endometriosis; skin
CC conditions such as dermatitis or urticaria; or immune system disorders
CC such as acquired immunodeficiency syndrome (AIDS). The nucleic acids may
CC additionally be used to generate animal models of cell death-associated
CC disorders. The present sequence represents a cell death protective
CC polypeptide
XX
XX Sequence 232 AA;
SQ
Query Match 14.8%; Score 89.5; DB 4; Length 232;
Best Local Similarity 29.3%; Pred. No. 0.42;
Matches 34; Conservative 15; Mismatches 60; Indels 7; Gaps 3;
OY 3 EQQRDLERKADTKNLERKKEHGDIILAEIDYGRLEIPIALPSENGEYIPIHQ--SSL 60
DB 72 EVKNEEEDQKDEEENQNEKAG--KEDDKKEDEGKEDKKNENKEDAKKEDEK 128
OY 61 PODNGNSRDSKEISIIKTNRESITNV--EGRRDIHKHLEKKDGSIKPEQKE 114
DB 129 GEDGKGNEDEGKEKEDKEDEDRKETGVGKENEDGKEDKDKGKGVKXDEKE 184

KW macular degeneration; hypertension; myocardial infarction;
KW atherosclerosis; respiratory disorder; asthma; transgenic animal;
KW chronic obstructive pulmonary disease; neoplastic condition; cancer;
KW benign tumour; anaemia; gastrointestinal disorder; gastritis;
KW ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;
KW glomerulonephritis; cystitis; endometriosis; endocrine disorder;
KW Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;
KW urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS.
XX
OS Homo sapiens.
XX
XX WO200145638-A2.
XX
XX 28-JUN-2001.
XX
XX 11-DEC-2000; 2000WO-US033547.
XX
XX 14-DEC-1999; 99US-00461697.
XX
XX (COGE-) COGENT NEUROSCIENCE INC.
XX
XX Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;
XX WPI; 2001-390297/41.
XX N-PSDB; AAH84226, AAH84229.
XX
XX Novel protective sequence polynucleotides and polypeptides, used to
PT identify modulators of their expression and activity, which are used in
PT to treat central nervous system conditions, diseases and disorders.
XX
XX Claim 1, Fig 9C, 325pp; English.
XX
XX Sequences AAH84132-AAH84370 represent human nucleic acid sequences which
CC protect against cell death (i.e., apoptosis or necrosis). Sequences
CC AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,
CC AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,
CC while the remaining nucleic acid sequences within the range given above
CC represent the open reading frames (ORFs) of these cDNA clones. Sequences
CC AAH98610-AAH98829 represent the polypeptides encoded by the cell death
CC protective ORFs. The cell death protective cDNA clones are able to
CC prevent, delay or reverse progression through the apoptotic or necrotic
CC pathways when injected into a cell predisposed to or undergoing cell
CC death. The cell death protective nucleic acids and polypeptides can be
CC used in the diagnosis and treatment of disorders associated with cell
CC death, and to screen for compounds which modulate their activity or
CC expression. Such modulators, preferably a small organic molecule, an
CC antibody, a ribozyme, or an antisense molecule, can also be used to treat
CC cell death-related diseases. Such diseases include those associated with
CC the central nervous system including psychiatric or neurological
CC disorders, especially ischemia-related conditions such as strokes, and
CC also includes neurodegenerative disorders such as Alzheimer's disease,
CC Huntington's disease, or Parkinson's disease. The modulators may also be
CC used to treat infections such as ischemic encephalopathy or cerebral infarction;
CC vascular diseases such as diabetic retinopathy or macular degeneration;
CC hypertension; myocardial infarction; atherosclerosis; respiratory
CC conditions such as asthma or chronic obstructive pulmonary disease;
CC neoplastic conditions such as cancers or benign tumours; blood cell
CC conditions such as anaemia; gastrointestinal conditions such as gastritis
CC or ulcerative colitis; liver conditions such as biliary cirrhosis; kidney
CC disorders such as glomerulonephritis; cystitis; endometriosis; endocrine
CC disorders such as Grave's disease or Hashimoto's thyroiditis; skin
CC conditions such as dermatitis or urticaria; or immune system disorders
CC such as acquired immunodeficiency syndrome (AIDS). The nucleic acids may
CC additionally be used to generate animal models of cell death-associated
CC disorders. The present sequence represents a cell death protective
CC polypeptide
XX
XX Sequence 238 AA;
SQ
Query Match 14.8%; Score 89.5; DB 4; Length 238;
Best Local Similarity 29.3%; Pred. No. 0.44;
Matches 34; Conservative 15; Mismatches 60; Indels 7; Gaps 3;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:16:04 ; Search time 11.6189 Seconds
(without alignments)
832.525 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LOEQQRDLQKRAADTKKLE.....GHLEKKQDSIKPEQKEDKS 117

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/prodata/1/1aa/5_COMB.pep.*
2: /cgn2_6/prodata/1/1aa/6_COMB.pep.*
3: /cgn2_6/prodata/1/1aa/H_COMB.pep.*
4: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep.*
5: /cgn2_6/prodata/1/1aa/RE_COMB.pep.*
6: /cgn2_6/prodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	603	100.0	117	2	US-08-098-327E-20	Sequence 20, Appl
2	603	100.0	117	2	US-08-462-625-20	Sequence 20, Appl
3	603	100.0	117	2	US-08-098-327E-31	Sequence 31, Appl
4	603	100.0	117	2	US-08-462-625-31	Sequence 31, Appl
5	554	91.9	107	2	US-08-098-327E-19	Sequence 19, Appl
6	554	91.9	107	2	US-08-462-625-19	Sequence 19, Appl
7	154	25.5	31	2	US-08-098-327E-23	Sequence 23, Appl
8	154	25.5	31	2	US-08-462-625-23	Sequence 23, Appl
9	141	23.4	27	2	US-08-098-327E-21	Sequence 21, Appl
10	141	23.4	27	2	US-08-462-625-21	Sequence 21, Appl
11	97	16.1	709	2	US-09-248-796A-19045	Sequence 19045, A
12	90.5	15.0	1085	1	US-08-431-080-28	Sequence 28, Appl
13	90.5	15.0	1085	1	US-08-938-534-28	Sequence 28, Appl
14	90.5	15.0	1085	2	US-09-345-284-28	Sequence 28, Appl
15	89.5	14.8	231	2	US-09-461-697-194	Sequence 194, App
16	89.5	14.8	231	2	US-09-461-697-192	Sequence 192, App
17	89.5	14.8	238	2	US-09-461-697-190	Sequence 190, App
18	89.5	14.8	257	2	US-09-461-697-188	Sequence 188, App
19	89.5	14.8	272	2	US-09-461-697-186	Sequence 186, App
20	86	14.3	447	2	US-09-710-279-2900	Sequence 2900, Ap
21	86	14.3	450	2	US-09-710-279-3226	Sequence 3226, Ap
22	85.5	14.2	765	2	US-09-538-092-906	Sequence 906, App
23	85.5	14.2	765	2	US-09-882-274-2	Sequence 2, Appl
24	85	14.1	15	2	US-09-109-953-74	Sequence 74, Appl
25	84.5	14.0	150	2	US-09-395-689-2	Sequence 2, Appl
26	84.5	14.0	300	2	US-09-395-689-1	Sequence 1, Appl
27	84.5	14.0	584	2	US-09-949-016-8587	Sequence 8587, Ap

28	84.5	14.0	585	2	US-09-949-016-6627	Sequence 6627, Ap
29	84.5	14.0	765	1	US-08-663-112-2	Sequence 2, Appl
30	84	13.9	24	2	US-08-098-327E-22	Sequence 22, Appl
31	84	13.9	24	2	US-08-462-625-22	Sequence 22, Appl
32	84	13.9	1761	2	US-09-561-709B-1	Sequence 1, Appl
33	83.5	13.8	1025	2	US-09-538-092-334	Sequence 334, App
34	82.5	13.7	1027	2	US-09-514-259-27	Sequence 27, Appl
35	80.5	13.3	129	2	US-09-513-999C-5707	Sequence 5707, Ap
36	80.5	13.3	346	2	US-09-902-540-12787	Sequence 12787, A
37	80.5	13.3	1032	2	US-09-914-259-26	Sequence 26, Appl
38	80.5	13.3	1032	2	US-09-538-092-1293	Sequence 1293, Ap
39	79	13.1	140	2	US-09-710-279-2268	Sequence 2268, Ap
40	79	13.1	140	2	US-09-710-279-2438	Sequence 2438, Ap
41	79	13.1	147	2	US-09-134-001C-5518	Sequence 5518, Ap
42	79	13.1	538	2	US-09-252-991A-23551	Sequence 23551, A
43	79	13.1	708	2	US-08-235-836C-76	Sequence 76, Appl
44	78	12.9	452	2	US-09-248-796A-14553	Sequence 14553, A
45	78	12.9	1270	2	US-09-538-092-1321	Sequence 1321, Ap

ALIGNMENTS

RESULT 1
US-08-098-327E-20
; Sequence 20, Application US/08098327E
; Patent No. 6270771
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098,327E
; FILING DATE: 24-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SRO ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; US-08-098-327E-20
Query Match 100.0%; Score 603; DB 2; Length 117;

Best Local Similarity 100.0%; Pred. No. 2,4e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LQEOQRDLQKADTKKLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 60
Db 1 LQEOQRDLQKADTKKLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 60
Qy 61 PDNRGNSRDSKESIIIEKTRESITTNVEGRDIHKHGLEKKDGSIKPEOKEDXS 117
Db 61 PDNRGNSRDSKESIIIEKTRESITTNVEGRDIHKHGLEKKDGSIKPEOKEDXS 117

RESULT 2

US-08-462-625-20
; Sequence 20, Application US/08462625
; Patent No. 6319502
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,625
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/098,327
; FILING DATE: 24-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; US-08-462-625-20

Query Match 100.0%; Score 603; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 2,4e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LQEOQRDLQKADTKKLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 60
Db 1 LQEOQRDLQKADTKKLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 60
Qy 61 PDNRGNSRDSKESIIIEKTRESITTNVEGRDIHKHGLEKKDGSIKPEOKEDXS 117

Db 61 PDNRGNSRDSKESIIIEKTRESITTNVEGRDIHKHGLEKKDGSIKPEOKEDXS 117

RESULT 3

US-08-098-327E-31
; Sequence 31, Application US/08098327E
; Patent No. 6270771
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098,327E
; FILING DATE: 24-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; US-08-098-327E-31

Query Match 100.0%; Score 603; DB 2; Length 316;
Best Local Similarity 100.0%; Pred. No. 9,4e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LQEOQRDLQKADTKKLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 60
Db 200 LQEOQRDLQKADTKKLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 259
Qy 61 PDNRGNSRDSKESIIIEKTRESITTNVEGRDIHKHGLEKKDGSIKPEOKEDXS 117
Db 260 PDNRGNSRDSKESIIIEKTRESITTNVEGRDIHKHGLEKKDGSIKPEOKEDXS 316

RESULT 4

US-08-462-625-31
; Sequence 31, Application US/08462625
; Patent No. 6319502
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUILHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-31
Query Match 100.0%; Score 603; DB 2; Length 316;
Best Local Similarity 100.0%; Pred. No. 9.4e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LQEQRDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALPSPENERGYIIPHOSSL 60
DB 200 LQEQRDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALPSPENERGYIIPHOSSL 259
QY 61 PQQNRGNSRDSKSIISTEKTNRRESITTVNGRRDINHGHLEKKDGSIRPEQKEDS 117
DB 260 PQQNRGNSRDSKSIISTEKTNRRESITTVNGRRDINHGHLEKKDGSIRPEQKEDS 316
RESULT 5
US-08-098-327E-19
Sequence 19, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria

STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-19
Query Match 91.9%; Score 554; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 8.4e-56;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RKADTKNLERKKEHGDILAEDLYGRLEIPALPSPENERGYIIPHOSSLPQNRNSRD 70
DB 1 RKADTKNLERKKEHGDILAEDLYGRLEIPALPSPENERGYIIPHOSSLPQNRNSRD 60
QY 71 SKSISTEKTNRRESITTVNGRRDINHGHLEKKDGSIRPEQKEDS 117
DB 61 SKSISTEKTNRRESITTVNGRRDINHGHLEKKDGSIRPEQKEDS 107
RESULT 6
US-08-462-625-19
Sequence 19, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-19

Query Match 91.9%; Score 554; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 8,4e-56;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADPKKLEKKEKGGDILAEEDLYGRLEIPALPEPSENERGYIIPQSSLPQDRNGNRD 70
Db 1 RKADPKKLEKKEKGGDILAEEDLYGRLEIPALPEPSENERGYIIPQSSLPQDRNGNRD 60

Qy 71 SKEISIIKTNRESITTVNVEGRDIIKHGHLEKKDGSIKPEQKEDKS 117
Db 61 SKEISIIKTNRESITTVNVEGRDIIKHGHLEKKDGSIKPEQKEDKS 107

RESULT 7
US-08-098-327E-23
Sequence 23, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-23

Query Match 25.5%; Score 154; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 9e-11;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 NSRDSKESIIKTNRESITTVNVEGRDIIKH 97
Db 1 NSRDSKESIIKTNRESITTVNVEGRDIIKH 31

RESULT 8
US-08-462-625-23
Sequence 23, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992
US-08-462-625-23

Query Match 25.5%; Score 154; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 9e-11;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 NSRDSKEISIIIEKTNRESITTVNVEGRDHIK 97
Db 1 NSRDSKEISIIIEKTNRESITTVNVEGRDHIK 31

RESULT 9
US-08-098-327E-21

Sequence 21, Application US/08098327E
Patent No. 6270771

GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine

APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE

TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404

CITY: Alexandria
STATE: Virginia

COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993

CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids

TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide

PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-21

Query Match 23.4%; Score 141; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 DTKKNLERKKEHGDIIAEDLYGRLEIP 40
Db 1 DTKKNLERKKEHGDIIAEDLYGRLEIP 27

RESULT 10
US-08-462-625-21

Sequence 21, Application US/08462625

Patent No. 6319502
GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE

TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404
CITY: Alexandria

STATE: Virginia
COUNTRY: United States

ZIP: 22313-1404
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625

FILING DATE: 05-JUN-1995
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327

FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids

TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide

PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992
US-08-462-625-21

Query Match 23.4%; Score 141; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 DTKKNLERKKEHGDIIAEDLYGRLEIP 40
Db 1 DTKKNLERKKEHGDIIAEDLYGRLEIP 27

RESULT 11
US-09-248-796A-19045

Sequence 19045, Application US/09248796A
Patent No. 6747137

GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196,132
CURRENT APPLICATION NUMBER: US/09/248,796A

PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 19045
LENGTH: 709
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-19045

Query Match 16.1%; Score 97; DB 2; Length 709;
Best Local Similarity 24.7%; Pred. No. 0.023;
Matches 37; Conservative 29; Mismatches 36; Indels 48; Gaps 9;

QY 11 RKADTKKLEKKEKHGDI-----LAEDLYGRLEIPA-----IELPSENERGYI 54
DB 462 KEVDTDLSIQKELFQIKEPVNDKDEKQIKEDLV-KVETPSPKDEGPAKVEVPSKDETPKXI 520
QY 55 PHQSSLPQDNNGNSRDSK-----EISIEKTNREST-----TTNVEGRRD 94
DB 521 QNIDS-TTND-GKAKDEKQNVSKVDSNVDTKDQDQIKVKEKPEKEDDDEGKCD 578
QY 95 -----IHKGHLEKKKDGSIKPE---OKED 115
DB 579 NEKVTVHKNEEDKDKGVNEAKDKVKQKD 608

RESULT 12

US-08-431-080-28
Sequence 28, Application US/08431080
Patent No. 3698686
GENERAL INFORMATION:
APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, Miriam S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,080
CLASSIFICATION: 514
FILING DATE: Concurrently Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-431-080-28

Query Match 15.0%; Score 90.5; DB 1; Length 1085;
Best Local Similarity 20.8%; Pred. No. 0.23;
Matches 31; Conservative 21; Mismatches 54; Indels 43; Gaps 4;

QY 3 EQRDLQRKADTKKLEKKE-----HGD 27
DB 865 DKMKEIQKTKTEKKRQLKKKKLLKIRKQKAIKEQETMNLQLGINGHEIIGNNNSHD 924
QY 28 I-LAEDLYGRLEIPAILEPSE-NEGYIIPHQSSLPQDNNGNSRDSKEI-----SIEK 79
DB 925 INTGDTFTNTENTPMNELPSHAPEDASLIPIHNSDLAVSNTRKSTKSVGLDEIHEILGK 984
QY 80 TNRSEITTVNTEGRDIIKHGLEEKDGSI 108
DB 985 DENDLSVGDINGIDAOEGHVIETPDADI 1013

RESULT 13

US-08-938-534-28
Sequence 28, Application US/08938534
Patent No. 5916752
GENERAL INFORMATION:
APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, Miriam S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,534
FILING DATE: 26-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,080
FILING DATE:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-938-534-28

Query Match 15.0%; Score 90.5; DB 1; Length 1085;
Best Local Similarity 20.8%; Pred. No. 0.23;
Matches 31; Conservative 21; Mismatches 54; Indels 43; Gaps 4;

QY 3 EQRDLQRKADTKKLEKKE-----HGD 27
DB 865 DKMKEIQKTKTEKKRQLKKKKLLKIRKQKAIKEQETMNLQLGINGHEIIGNNNSHD 924
QY 28 I-LAEDLYGRLEIPAILEPSE-NEGYIIPHQSSLPQDNNGNSRDSKEI-----SIEK 79
DB 925 INTGDTFTNTENTPMNELPSHAPEDASLIPIHNSDLAVSNTRKSTKSVGLDEIHEILGK 984
QY 80 TNRSEITTVNTEGRDIIKHGLEEKDGSI 108

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:18:34 ; Search time 30.5335 Seconds
(without alignments)
1601.063 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LDEQQRDLERKADTKKLE.....GHLEKKDSIKPEQKEDKS 117

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	603	100.0	117	3	US-09-837-344-20
2	603	100.0	117	3	US-09-900-963-20
3	603	100.0	117	3	US-09-837-344-31
4	603	100.0	316	3	US-09-900-963-31
5	602	99.8	1909	5	US-10-732-923-3341
6	554	91.9	107	3	US-09-837-344-19
7	554	91.9	107	3	US-09-900-963-19
8	154	25.5	31	3	US-09-837-344-23
9	154	25.5	31	3	US-09-900-963-23
10	141	23.4	27	3	US-09-837-344-21
11	141	23.4	27	3	US-09-900-963-21
12	94	15.6	456	3	US-09-894-018-121
13	94	15.6	456	3	US-10-474-960A-121
14	90	14.9	756	5	US-10-756-149-5156
15	89.5	14.8	231	3	US-09-922-261-194
16	89.5	14.8	232	3	US-09-922-261-192
17	89.5	14.8	238	3	US-09-922-261-190
18	89.5	14.8	257	3	US-09-922-261-188
19	89.5	14.8	272	3	US-09-922-261-186
20	87	14.4	550	4	US-10-369-493-5995
21	85.5	14.2	765	4	US-09-882-274-2
22	85.5	14.2	765	4	US-10-408-765A-1149
23	85.5	14.2	765	4	US-10-484-577-679
24	85	14.1	15	4	US-10-103-395-74
25	84.5	14.0	242	5	US-10-450-763-32854
26	84.5	14.0	585	5	US-10-631-467-870
27	84	13.9	24	3	US-09-837-344-22

28	84	13.9	24	3	US-09-900-963-22	Sequence 22, Appl
29	83.5	13.8	257	4	US-10-425-114-58860	Sequence 58860, A
30	83.5	13.8	395	4	US-10-425-114-69744	Sequence 69744, A
31	83.5	13.8	399	4	US-10-425-115-226794	Sequence 226794, A
32	83.5	13.8	419	4	US-10-425-115-226797	Sequence 226797, A
33	83.5	13.8	807	5	US-10-511-698-39	Sequence 39, Appl
34	83.5	13.8	903	4	US-10-282-122A-52328	Sequence 52328, A
35	83	13.8	2683	4	US-10-437-963-185370	Sequence 185370, A
36	82.5	13.7	1027	4	US-10-080-608A-27	Sequence 27, Appl
37	82.5	13.7	127	4	US-10-370-685-116	Sequence 116, Appl
38	82.5	13.7	1264	4	US-10-369-493-3742	Sequence 3742, Ap
39	82.5	13.7	1642	5	US-10-732-923-8639	Sequence 8639, Ap
40	82	13.6	472	3	US-09-953-407-2	Sequence 2, Appl
41	82	13.6	1429	5	US-09-953-407-1	Sequence 1, Appl
42	82	13.6	1429	5	US-10-450-763-35896	Sequence 35896, A
43	81.5	13.5	485	3	US-09-769-736-72	Sequence 72, Appl
44	81.5	13.5	980	5	US-10-732-923-3343	Sequence 3343, Ap
45	81.5	13.5	2447	5	US-10-450-763-52739	Sequence 52739, A

ALIGNMENTS

RESULT 1
US-09-837-344-20
; Sequence 20, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-09-837-344-20

Query Match 100.0%; Score 603; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 6.8e-51;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LOEQRDLQKADTKKLNLERKKEHGDILAEGLYGRLEIPALELPSENERGYIIPHOSSL 60
Db 1 LOEQRDLQKADTKKLNLERKKEHGDILAEGLYGRLEIPALELPSENERGYIIPHOSSL 60
Qy 61 PQDRNGSRDSKESIIIEKTNRESITTVNEGRDIIHGKLEKKDGSIKPEQKEDKS 117
Db 61 PQDRNGSRDSKESIIIEKTNRESITTVNEGRDIIHGKLEKKDGSIKPEQKEDKS 117

RESULT 2

US-09-900-963-20
; Sequence 20, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine
DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSER: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22133-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/900,963

FILING DATE: 10-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/098,327

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-045

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 117 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992

SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-09-900-963-20

Query Match 100.0%; Score 603; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 6.8e-51;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LOEQRDLQKADTKKLNLERKKEHGDILAEGLYGRLEIPALELPSENERGYIIPHOSSL 60
Db 1 LOEQRDLQKADTKKLNLERKKEHGDILAEGLYGRLEIPALELPSENERGYIIPHOSSL 60

Qy 61 PQDRNGSRDSKESIIIEKTNRESITTVNEGRDIIHGKLEKKDGSIKPEQKEDKS 117
Db 61 PQDRNGSRDSKESIIIEKTNRESITTVNEGRDIIHGKLEKKDGSIKPEQKEDKS 117

RESULT 3

US-09-837-344-31
; Sequence 31, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine
DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSER: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22133-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/837,344

FILING DATE: 19-Apr-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/462,625

FILING DATE: <Unknown>

APPLICATION NUMBER: FR 91 01286

FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-078

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 316 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992

SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-09-837-344-31

Query Match 100.0%; Score 603; DB 3; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.4e-50;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LOEQRDLQKADTKKLNLERKKEHGDILAEGLYGRLEIPALELPSENERGYIIPHOSSL 60
Db 200 LOEQRDLQKADTKKLNLERKKEHGDILAEGLYGRLEIPALELPSENERGYIIPHOSSL 259
Qy 61 PQDRNGSRDSKESIIIEKTNRESITTVNEGRDIIHGKLEKKDGSIKPEQKEDKS 117
Db 260 PQDRNGSRDSKESIIIEKTNRESITTVNEGRDIIHGKLEKKDGSIKPEQKEDKS 316

RESULT 4

US-09-900-963-31
; Sequence 31, Application US/09900963

```
/ Publication No. US20030064075A1
/ GENERAL INFORMATION:
/ APPLICANT: GUERIN-MARCHAND, Claudine
/ DRUILHE, Pierre
/ TITLE OF INVENTION: HEPATIC SEQUENCES SPECIFIC FOR THE
/ HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
/ OF STIMULATING THE T LYMPHOCYTES
/ NUMBER OF SEQUENCES: 46
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Burns, Doane, Swecker & Mathis
/ STREET: P.O. Box 1404
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: United States
/ ZIP: 22313-1404
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ SOFTWARE: Patentin Releasee #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/900,963
/ FILING DATE: 10-Jul-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/098,327
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McGowan, Malcolm K.
/ REGISTRATION NUMBER: 39,300
/ REFERENCE/DOCKET NUMBER: 010830-045
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-2021
/ INFORMATION FOR SEQ ID NO: 31:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 316 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ PUBLICATION INFORMATION:
/ DOCUMENT NUMBER: WO 92/13884
/ PUBLICATION DATE: 20-AUG-1992
/ SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-900-963-31
Query Match 100.0%; Score 603; DB 3; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.4e-50;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LOEQORDLEQRKADTKNLERKKEHGDILAE DLYGRLEIPALILPSENERGYIIPHOSSL 60
DB 200 LOEQORDLEQRKADTKNLERKKEHGDILAE DLYGRLEIPALILPSENERGYIIPHOSSL 259
QY 61 PDNRGNSRDSKEISIIKTNRESITTVNVEGRDINHKGHLKKDGSIKPEQKEDKS 117
DB 260 PDNRGNSRDSKEISIIKTNRESITTVNVEGRDINHKGHLKKDGSIKPEQKEDKS 316
RESULT 5
US-10-732-923-3341
/ Sequence 3341, Application US/10732923
/ Publication No. US20050108791A1
/ GENERAL INFORMATION:
/ APPLICANT: Edgerton, Michael D
/ TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
/ FILE REFERENCE: 38-15(52796)C
/ CURRENT APPLICATION NUMBER: US/10/732,923
/ CURRENT FILING DATE: 2003-12-10
/ PRIOR APPLICATION NUMBER: 10/310,154
/ PRIOR FILING DATE: 2002-12-04
/ NUMBER OF SEQ ID NOS: 24149
```

```
/ SEQ ID NO 3341
/ LENGTH: 1909
/ TYPE: PRT
/ ORGANISM: Plasmodium falciparum
/ US-10-732-923-3341
Query Match 99.8%; Score 602; DB 5; Length 1909;
Best Local Similarity 99.1%; Pred. No. 2.9e-49;
Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LOEQORDLEQRKADTKNLERKKEHGDILAE DLYGRLEIPALILPSENERGYIIPHOSSL 60
DB 1620 LOEQORDLEQRKADTKNLERKKEHGDILAE DLYGRLEIPALILPSENERGYIIPHOSSL 1679
QY 61 PDNRGNSRDSKEISIIKTNRESITTVNVEGRDINHKGHLKKDGSIKPEQKEDKS 117
DB 1680 PDNRGNSRDSKEISIIKTNRESITTVNVEGRDINHKGHLKKDGSIKPEQKEDKS 1736
RESULT 6
US-09-837-344-19
/ Sequence 19, Application US/09837344
/ Patent No. US20020041882A1
/ GENERAL INFORMATION:
/ APPLICANT: GUERIN-MARCHAND, Claudine
/ DRUILHE, Pierre
/ TITLE OF INVENTION: HEPATIC SEQUENCES SPECIFIC FOR THE
/ HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
/ OF STIMULATING THE T LYMPHOCYTES
/ NUMBER OF SEQUENCES: 46
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Burns, Doane, Swecker & Mathis
/ STREET: P.O. Box 1404
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: United States
/ ZIP: 22313-1404
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ SOFTWARE: Patentin Releasee #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/837,344
/ FILING DATE: 19-Apr-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/462,625
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: FR 91 01286
/ FILING DATE: 05-FEB-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McGowan, Malcolm K.
/ REGISTRATION NUMBER: 39,300
/ REFERENCE/DOCKET NUMBER: 010830-078
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-2021
/ INFORMATION FOR SEQ ID NO: 19:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 107 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ PUBLICATION INFORMATION:
/ DOCUMENT NUMBER: WO 92/13884
/ PUBLICATION DATE: 20-AUG-1992
/ SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-837-344-19
Query Match 91.9%; Score 554; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 3.6e-46;
```

Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADTKNLERKKEKGDILADLYGRLEIPALELPSENERGYITPHOSLPODNRGNSRD 70
Db 1 RKADTKNLERKKEKGDILADLYGRLEIPALELPSENERGYITPHOSLPODNRGNSRD 60
Qy 71 SKEISIEKTNRESITTVNVEGRDIIKHGLEEKDGSIKPKOKEDKS 117
Db 61 SKEISIEKTNRESITTVNVEGRDIIKHGLEEKDGSIKPKOKEDKS 107

RESULT 7

US-09-900-963-19
; Sequence 19, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,963
; FILING DATE: 10-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/098,327
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-900-963-19

Query Match 91.9%; Score 554; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 3.6e-46;

Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADTKNLERKKEKGDILADLYGRLEIPALELPSENERGYITPHOSLPODNRGNSRD 70
Db 1 RKADTKNLERKKEKGDILADLYGRLEIPALELPSENERGYITPHOSLPODNRGNSRD 60
Qy 71 SKEISIEKTNRESITTVNVEGRDIIKHGLEEKDGSIKPKOKEDKS 117
Db 61 SKEISIEKTNRESITTVNVEGRDIIKHGLEEKDGSIKPKOKEDKS 107

RESULT 8

US-09-837-344-23
; Sequence 23, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-837-344-23

Query Match 25.5%; Score 154; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 6.6e-08;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 NSRDSKESIEKTNRESITTVNVEGRDIIKH 97
Db 1 NSRDSKESIEKTNRESITTVNVEGRDIIKH 31

RESULT 9

US-09-900-963-23
; Sequence 23, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46

;;
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burns, Doane, Swecker & Mathis
;; STREET: P.O. Box 1404
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/900,963
;; FILING DATE: 10-Jul-2001
;; CLASSIFICATION: <Unknown>
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/098,327
;; FILING DATE: <Unknown>
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McGowan, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 010830-045
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2021
;;
;; INFORMATION FOR SEQ ID NO: 23:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 31 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; PUBLICATION INFORMATION:
;; DOCUMENT NUMBER: WO 92/13884
;; PUBLICATION DATE: 20-AUG-1992
;; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-900-963-23
;
Query Match 25.5%; Score 154; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 6,6e-08;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 67 NSRDSKSIIEKTNRESITTVNGRRDIHK 97
DB 1 NSRDSKSIIEKTNRESITTVNGRRDIHK 31
;
RESULT 10
US-09-837-344-21
; Sequence 21, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344

;;
;; FILING DATE: 19-Apr-2001
;; CLASSIFICATION: <Unknown>
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/462,625
;; FILING DATE: <Unknown>
;; APPLICATION NUMBER: FR 91 01286
;; FILING DATE: 05-FEB-1991
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McGowan, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 010830-078
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2021
;;
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 27 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; PUBLICATION INFORMATION:
;; DOCUMENT NUMBER: WO 92/13884
;; PUBLICATION DATE: 20-AUG-1992
;; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-837-344-21
;
Query Match 23.4%; Score 141; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 14 DTKNLERKKEHGDILAEIDLXGRLEIP 40
DB 1 DTKNLERKKEHGDILAEIDLXGRLEIP 27
;
RESULT 11
US-09-900-963-21
; Sequence 21, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,963
; FILING DATE: 10-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/098,327
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021


```

; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Piranaw, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/922,261
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US/09/461,697
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 194
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-261-194

```

```

Query Match      14.8%; Score 89.5; DB 3; Length 231;
Best Local Similarity 29.3%; Pred. No. 1.6;
Matches 34; Conservative 15; Mismatches 60; Indels 7; Gaps 3;

QY      3  EGOQDLQKAKDTKKNLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIYPHQ--SSL 60
DB      71  EVKNEEDQKEDKEDQNEKEGANG--KEDKDEKGEEDGKEDKNGNEKGEADAKEDGKK 127

QY      61  PODNRGNSRDSKESISIEKTNRESITTV--EGRRDIHKHGLEKKDGSIKPEQKE 114
DB      128  GEDGKNGEDGKEGEDEKEDRKEKTVGKGNEDGKKGDKKGDVVKYKEDDEKE 183

```

Search completed: December 7, 2005, 07:30:34
 Job time : 31.5335 sec

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 7, 2005, 07:21:30, Search time 2.70208 Seconds
(without alignments)
241.806 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LOEQORDLEFORKADTKKNE.....GHLEKKKGSIKPEKEDKS 117

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 32527 seqs, 5584426 residues

Total number of hits satisfying chosen parameters: 32527

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA_New.*
1: /cgn2_6/prodata/2/pubppa/US09_NEW_PUB pep.*
2: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB pep.*
3: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB pep.*
4: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB pep.*
5: /cgn2_6/prodata/2/pubppa/PCF_NEW_PUB pep.*
6: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB pep.*
7: /cgn2_6/prodata/2/pubppa/US11_NEW_PUB pep.*
8: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	14.3	447	US-10-793-626-2900	Sequence 2900, Ap
2	86	14.3	450	US-10-793-626-3226	Sequence 3226, Ap
3	79.5	13.2	2432	US-10-821-234-899	Sequence 899, App
4	79	13.1	140	US-10-793-626-2268	Sequence 2268, Ap
5	79	13.1	140	US-10-793-626-2438	Sequence 2438, Ap
6	79	13.1	795	US-10-770-726-49	Sequence 49, App1
7	76.5	12.7	677	US-10-982-545-12	Sequence 12, App1
8	75	12.4	485	US-10-508-263-22	Sequence 22, App1
9	72.5	12.0	491	US-10-793-626-2808	Sequence 2808, Ap
10	71.5	11.9	212	US-10-467-657-5486	Sequence 5486, Ap
11	71.5	11.9	212	US-10-467-657-7404	Sequence 7404, Ap
12	71	11.8	328	US-10-467-657-7710	Sequence 7710, Ap
13	69	11.4	793	US-11-060-914-2	Sequence 2, App11
14	68	11.3	531	US-11-060-914-4	Sequence 4, App11
15	67	11.1	369	US-10-467-657-1336	Sequence 1336, Ap
16	67	11.1	904	US-10-507-275-5	Sequence 5, App11
17	66	10.9	495	US-10-467-657-3844	Sequence 3844, Ap
18	66	10.9	658	US-10-467-657-4782	Sequence 4782, Ap
19	66	10.9	676	US-11-135-855-28	Sequence 28, App1
20	66	10.9	717	US-11-135-855-29	Sequence 29, App1
21	66	10.9	862	US-11-077-550-171	Sequence 171, App
22	66	10.9	862	US-11-077-550-173	Sequence 173, App
23	65.5	10.9	257	US-10-793-626-126	Sequence 126, App
24	65.5	10.9	397	US-10-793-626-1244	Sequence 1244, Ap
25	65.5	10.9	578	US-10-821-234-1039	Sequence 1039, Ap

26	65	10.8	360	US-10-878-556A-134	Sequence 134, App
27	65	10.8	360	US-11-186-284-115	Sequence 115, App
28	65	10.8	790	US-10-763-712A-75	Sequence 75, App1
29	64.5	10.7	684	US-10-793-626-2098	Sequence 2098, Ap
30	64.5	10.7	1142	US-11-109-156-22	Sequence 22, App1
31	64	10.6	904	US-11-087-227-12	Sequence 12, App1
32	63.5	10.5	215	US-10-793-626-1650	Sequence 1650, Ap
33	63.5	10.5	286	US-10-793-626-2192	Sequence 2192, Ap
34	63.5	10.5	612	US-10-467-657-3988	Sequence 3988, Ap
35	63.5	10.5	664	US-10-878-556A-44	Sequence 44, App1
36	63	10.4	963	US-10-467-962B-2	Sequence 2, App11
37	63	10.4	1404	US-10-878-556A-169	Sequence 169, App
38	62.5	10.4	505	US-10-821-234-1287	Sequence 1287, Ap
39	62	10.3	548	US-11-077-550-24	Sequence 24, App1
40	62	10.3	858	US-11-077-550-22	Sequence 22, App1
41	62	10.3	1169	US-11-077-550-20	Sequence 20, App1
42	61.5	10.2	387	US-10-485-517-132	Sequence 132, App
43	61.5	10.2	451	US-10-467-657-366	Sequence 366, App
44	61.5	10.2	635	US-10-821-234-1573	Sequence 1573, Ap
45	61.5	10.2	858	US-10-645-441-25	Sequence 25, App1

ALIGNMENTS

RESULT 1
US-10-793-626-2900
; Sequence 2900, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIORITY FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2900
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2900

Query Match 14.3%; Score 86; DB 6; Length 447;
Best Local Similarity 23.8%; Pred. No. 0.18; Mismatches 26; Gaps 4;
Matches 29; Conservative 24; Indels 43; Gaps 4;
QY 1 LOEQORDLEFORKADTKKNE-----RKKEHGDIADLYGRLETPALEPSENERGYYP 55
DB 15 ISKAQAVAEKQDLETFANLKADIDAKKEYEL--EQLSKEIEASA-----P 59
QY 56 HGGSLPDQNRGNSRDSKEISIEIKTNRESITTVNBSGRDIIHKHGLEKKDGSIKPEKED 115
DB 60 KQDEPPKDGAEVEDNKGSGSESNKPSDDEPEGTSD-----BEKPDAPKPDKPE 113
QY 116 KS 117
DB 114 ET 115
RESULT 2
US-10-793-626-3226
; Sequence 3226, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US


```

RESULT 6
US-10-770-726-49
; Sequence 49, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: CANCERS
; FILE REFERENCE: AM101079 (031896-010000)
; CURRENT APPLICATION NUMBER: US/10/770,726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 795
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-770-726-49

```

Query Match	13.1%	Score 79;	DB 6;	Length 795;
Best Local Similarity	28.2%	Pred. No. 1.7;		
Matches	35;	Conservative	20;	Mismatches 49;
			Indels	20;
			Gaps	7;

```
Oy      1 LQEQQRPLEQ--RADTK--KNLEKKKEHGDI LAEDLYGRLEIPALELPSENGEYIIPHQ   57
       |||::|||:|||:::|||||:::|||||
Db     17 LQEKRRKEQEKAETIKLKLNSDDRSKDSL EE---GELRDHREMITRNS-----PYR   68
```

QY 58 GSLPQNRGNRSDS-----KSTIIETNRRESITTVNVEGRRIHKGH-LEEKDGSIKP 110

Db 69 REDSMEDRGEEEDSLAKPQOMSRKKEVHHHRKDEKREKRR--HRSLSAEGGCHARYKE 126

Qy	111	EQKE	114
		:::	
Db	127	KERE	130

RESULT 7
US-10-982-545-12
; Sequence 12, Application US/10982545
; Publication No. US20050244890A1

APPLICANT: Davies, Huw Alun
APPLICANT: McGuire, James
APPLICANT: Simonsen, Anja Hvid
APPLICANT: Blomqvist, Kaj
APPLICANT: Podusz, Vladimir
APPLICANT: Ciplergen Biosystems, Inc.
TITLE OF INVENTION: Biomarkers For Alzheimer's Disease
FILE REFERENCE: 01866-0115505
CURRENT APPLICATION NUMBER: US/10/982,545

```

/ PRIOR APPLICATION NUMBER: US 60/526,753
/ PRIOR FILING DATE: 2003-12-02
/ PRIOR APPLICATION NUMBER: US 60/546,423
/ PRIOR FILING DATE: 2004-02-19
/ PRIOR APPLICATION NUMBER: US 60/547,250
/ PRIOR FILING DATE: 2004-02-23
/ PRIOR APPLICATION NUMBER: US 60/558,896
/ PRIOR FILING DATE: 2004-04-02
/ PRIOR APPLICATION NUMBER: US 60/572,617
/ PRIOR FILING DATE: 2004-05-18
/ PRIOR APPLICATION NUMBER: US 60/586,503
/ PRIOR FILING DATE: 2004-07-08
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 12
/ LENGTH: 677

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Chromogranin B precursor
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(20)
; OTHER INFORMATION: signal peptide
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (21)..(677)
; OTHER INFORMATION: Chromogranin B (Secretogranin I)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (326)..(385)
; OTHER INFORMATION: biomarker peptide 7258 Da, processed fragment of
; OTHER INFORMATION: Chromogranin B
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (341)
; OTHER INFORMATION: Xaa = sulfoltyrosine
US-10-982-545-12

```

Query Match	12.7%;	Score 76.5;	DB 6;	Length 677;
Best Local Similarity	23.2%;	Pred. No. 2.4;		
Matches	29;	Conservative	19;	Mismatches 58;
			Indels	19;
			Gaps	3;

```

Oy      4  QQRDIQRKADTKKXLERKKEHGDILAEIDYGRLEI PAI -----ELPSENERGYI  54
          ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      184 EKHLIEPGETQNAFLNERKQASAIKKEELVARSETHAAGSHQSEKTHSRKRSQSEGEFA  243

```

QY 55 PHSLLPNDNRGNSDSKESITIEKNRESIITTVNGRRDINKHGL-----EEKDGSI 108
 | | | | | | | | | | | | | | |
Db 244 GSQENHPQESKGPRSGEE---SEGEEDATSEVDKRRFTTPRNHNIGRSRPDRSSGGSL 299

QY	109	KPEQK	113
		:	
Db	300	PSEEK	304

```

RESULT 8
US-10-508-263-22
; Sequence 22, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085_00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 485
; TYPE: prt
; ORGANISM: Glycine max
US-10-508-263-22

```

[illegible]

```
RESULT 9
US-10-793-626-2808
; Sequence 2808, Application US/107933626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P034800S
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2808
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2808

Query Match
Best Local Similarity 21.9%; Score 72.5; DB 6; Length 491;
Matches 23; Conservative 24; Mismatches 37; Indels 21; Gaps 6;

Oy 7 DLEQKADTKKXLERKEHGDILAEGLRLEIPALPESENGYIIPHOSLPOD--N 64
Db 363 DQLRNVDISGLYLNKQSMVD---SWYSLSVF-----RGDYGFF-NRAYLPDALN 412
Oy 65 RGNRSRKEISITIEKTN-----RESITNV--EGRRDIHKAL 100
Db 413 KGNVNTKYVTQLIKELNTTFGDKQGVYNEILNESKIDIPSYI 457

RESULT 10
US-10-467-657-5486
; Sequence 5486, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5486
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5486

Query Match
Best Local Similarity 23.2%; Score 71.5; DB 6; Length 212;
Matches 35; Conservative 21; Mismatches 38; Indels 57; Gaps 8;

Oy 14 DTKNLERKKE-----HGDLA--EDLYGLEIPAI 42
Db 56 DSKLSEKRDALAEIKQAVAMHVAASTPEEIASLNIHTATLAKRAVYG----- 108
Oy 43 ELPSNER-----GYIIPHOSLPOD--RGNSR--DSKEISITIEKTNRESITTVNVEGR- 93
Db 109 -LAARPEKFIIDGNRIPEHIGIPAEAVVVGDSKTIIEISAAVLAKTARDAEMVYLAQRPP 167
```

```
Oy 94 ----DIHKG-----HLEKKDGSIKPEOKED 115
Db 168 QYGFDKHKGYGTQKHLEALKQYGVLPBHRRD 198

RESULT 11
US-10-467-657-7404
; Sequence 7404, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7404
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7404

Query Match
Best Local Similarity 22.2%; Score 71.5; DB 6; Length 212;
Matches 35; Conservative 21; Mismatches 38; Indels 57; Gaps 8;

Oy 14 DTKNLERKKE-----HGDLA--EDLYGLEIPAI 42
Db 56 DSKLSEKRDALAEIKQAVAMHVAASTPEEIASLNIHTATLAKRAVYG----- 108
Oy 43 ELPSNER-----GYIIPHOSLPOD--RGNSR--DSKEISITIEKTNRESITTVNVEGR- 93
Db 109 -LAARPEKFIIDGNRIPEHIGIPAEAVVVGDSKTIIEISAAVLAKTARDAEMVYLAQRPP 167
Oy 94 ----DIHKG-----HLEKKDGSIKPEOKED 115
Db 168 QYGFDKHKGYGTQKHLEALKQYGVLPBHRRD 198

RESULT 12
US-10-467-657-7710
; Sequence 7710, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7710
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7710

Query Match
Best Local Similarity 30.9%; Score 71; DB 6; Length 328;
Matches 35; Conservative 21; Mismatches 38; Indels 57; Gaps 8;
```


Matches 25; Conservative 15; Mismatches 35; Indels 6; Gaps 3;
 QY 35 GLEIPALPSENERGY-IPHOSSLPODNR--GNSRDSKISIIETKTNRESITTTNVG 91
 Db 137 GQIE---MEIKYEGGQGVSGRGVYVDENRQIGAIQDASFSPIRSVFEVAPVQ 193
 QY 92 RRDIIKHGLEEKDGSIKPEQ 112
 Db 194 RTDLKLVLDIETDGSIDPEE 214

RESULT 13
 US-11-060-914-2
 / Sequence 2, Application US/11060914
 / Publication No. US20050261184A1
 / GENERAL INFORMATION:
 / APPLICANT: Kaufman, Paul L
 / APPLICANT: Kaufman, Benjamin
 / APPLICANT: Berzhadsky, Alexander
 / APPLICANT: Borras, Teresa
 / TITLE OF INVENTION: Method for Treating Glaucoma
 / FILE REFERENCE: 960296.00128
 / CURRENT APPLICATION NUMBER: US/11/060,914
 / PRIOR FILING DATE: 2005-02-18
 / PRIOR APPLICATION NUMBER: US 60/545,722
 / PRIOR FILING DATE: 2004-02-18
 / PRIOR APPLICATION NUMBER: US 60/545,723
 / PRIOR FILING DATE: 2004-02-18
 / NUMBER OF SEQ ID NOS: 14
 / SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 2
 / LENGTH: 793
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 US-11-060-914-2

Query Match 11.4%; Score 69; DB 7; Length 793;
 Best Local Similarity 22.4%; Pred. No. 15;
 Matches 28; Conservative 34; Mismatches 39; Indels 24; Gaps 6;
 QY 1 LOEQORDLEQKAD-----TKNLERKKEHGDILAEDLYGRLEIPALPSENE 49
 Db 579 LEEERRRKQEEADRLREBEERKRLKEIERRR-----AEAAEKRGKMPBGL--SDCK 61
 QY 50 RGY--YIPHOSSLPODNRGN--SRDSKEISIIETKTNRESITTTNVGGRDIIKHGLEEKD 105
 Db 632 KPPKCTPFGSSSLKIERAFPLKSVQKSGVSTHQAIVSKIDRLLEGYTSALIGTK- 690
 QY 106 GSIXP 110
 Db 691 -SAKP 694

RESULT 14
 US-11-060-914-4
 / Sequence 4, Application US/11060914
 / Publication No. US20050261184A1
 / GENERAL INFORMATION:
 / APPLICANT: Kaufman, Paul L
 / APPLICANT: Geiger, Benjamin
 / APPLICANT: Berzhadsky, Alexander
 / APPLICANT: Borras, Teresa
 / TITLE OF INVENTION: Method for Treating Glaucoma
 / FILE REFERENCE: 960296.00128
 / CURRENT APPLICATION NUMBER: US/11/060,914
 / PRIOR FILING DATE: 2005-02-18
 / PRIOR APPLICATION NUMBER: US 60/545,722
 / PRIOR FILING DATE: 2004-02-18
 / PRIOR APPLICATION NUMBER: US 60/545,723
 / PRIOR FILING DATE: 2004-02-18
 / NUMBER OF SEQ ID NOS: 14
 / SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 4

/ LENGTH: 531
 / TYPE: PRT
 / ORGANISM: Rattus norvegicus
 US-11-060-914-4
 Query Match 11.3%; Score 68; DB 7; Length 531;
 Best Local Similarity 22.0%; Pred. No. 11;
 Matches 29; Conservative 31; Mismatches 44; Indels 28; Gaps 6;
 QY 1 LOEQORDLEQKADTKKLERKKEHGDILAEDLYGRLEIPALPSENERGYIPIHQ 57
 Db 91 LERLARREBRQKRLQEALEERQKEDPTITD---GSLVPSRMONNSAENE----- 139
 QY 58 SSLPODNRGNR---DSKEISIIETKTNRESITTTNVGGRDIIKHGLEEK-----D 105
 Db 140 -TAEGEKGESRSGGYEMETEVITTSYQKSYQDAEDKKEKEEEREBEELKGNLGE 198
 QY 106 GSIXPEQ-KEDK 116
 Db 199 NOIKDEKIKDK 210

RESULT 15
 US-10-467-657-1336
 / Sequence 1336, Application US/10467657
 / Publication No. US20050260581A1
 / GENERAL INFORMATION:
 / APPLICANT: CHIRON SPA
 / APPLICANT: FONTANA Maria Rita
 / APPLICANT: PIZZA Mariagrazia
 / APPLICANT: MANSIGNI Vega
 / APPLICANT: MONACI Elisabetta
 / TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 / FILE REFERENCE:
 / CURRENT APPLICATION NUMBER: US/10/467,657
 / PRIOR FILING DATE: 2003-08-11
 / PRIOR APPLICATION NUMBER: GB-0103424.8
 / PRIOR FILING DATE: 2001-02-12
 / NUMBER OF SEQ ID NOS: 9218
 / SOFTWARE: SeqMan99, version 1.04
 / SEQ ID NO 1336
 / LENGTH: 369
 / TYPE: PRT
 / ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-1336

Query Match 11.1%; Score 67; DB 6; Length 369;
 Best Local Similarity 27.4%; Pred. No. 9;
 Matches 20; Conservative 14; Mismatches 29; Indels 10; Gaps 1;
 QY 40 PALPSENERGYIIPHOSSLPODNRGNRD-----SKEISIIETKTNRESITTTNV 89
 Db 163 PGHTVSEYKRGGCVSLIAVYQDNGKAKDIALSYAANGGTGKGVIEETTFRETTDL 222
 QY 90 EGRDIIKHGLEE 102
 Db 223 FGEQAVLCGVAE 235

Search completed: December 7, 2005, 07:31:06
 Job time : 2.70208 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM proteinh - nucleic search, using frame_plus_pzn model

Run on: December 7, 2005, 07:30:46 ; Search time 2073.65 Seconds
(without alignments)
2639.584 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603

Sequence: 1 LQEQQRDLERKADTKKLE.....GHLEKKDSIKREKEDKS 117

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 8215650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ pzn.model -DEV=xlh
-Q/cgnt2.1/USPTO.spool h/US09837344/runat 05122005 133532 750/app query.fasta_1.718
-DB=EST -OPMT=fastap -SUPFIX=rat -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -STAR=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pcc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09837344 @CGN 1 1 6731 @runat 05122005 133532 750 -NCPU=6 -ICPU=3
-NO MMAP -JARGOUDRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	396	65.7	317	10	T09652 0224m3 gmbp
2	103.5	17.2	1782	2	BF133651 601778550
3	99.5	16.5	2131	10	AG332417 Mus muscu
4	98	16.3	589	7	CV271474 WS0154.B2
5	97.5	16.2	605	6	CA082727 SCBPAM202
6	96	15.9	684	8	DR983403 JGI AOKFI
7	96	15.9	1042	3	BI559491 603252811

8	96	15.9	1082	2	BG497835	BG497835 602543125
9	95	15.8	1714	10	AG390929	AG390929 Mus muscu
10	94.5	15.7	667	6	CA083387	CA083387 SCBPAM201
11	94.5	15.7	676	6	CA160204	CA160204 SCJFR23C0
12	94.5	15.7	720	8	CX328033	CX328033 JGI_X2T67
13	94.5	15.7	974	6	CA234102	CA234102 SCBPAM225
14	94.5	15.6	1665	2	BG284329	BG284329 602408426
15	94	15.6	691	2	BG832038	BG832038 602765038
16	94	15.6	920	5	BH132586	BH132586 ENFOQ42TR
17	94	15.6	1502	9	BU907774	BU907774 AGENCOURT
18	93	15.4	1218	10	CL080889	CL080889 CH216-160
19	92.5	15.3	598	10	CE688248	CE688248 tigr-gss-
20	92.5	15.3	621	10	CM816235	CM816235 pinea.259
21	92.5	15.3	1118	3	BO073445	BO073445 AGENCOURT
22	92.5	15.3	1595	10	AG365410	AG365410 Mus muscu
23	92	15.3	575	3	BM110778	BM110778 EST558314
24	92	15.3	647	5	BU295065	BU295065 603741656
25	92	15.3	959	2	B1114982	B1114982 602862806
26	91.5	15.2	386	9	AZ164328	AZ164328 SP_0075_B
27	91.5	15.2	663	3	B1068434	B1068434 B1068434
28	91.5	15.2	803	3	BP704622	BP704622 BP704622
29	91.5	15.2	1211	2	BF037585	BF037585 601461152
30	91.5	15.2	1340	10	AG334065	AG334065 Mus muscu
31	91.5	15.2	1513	10	AG453925	AG453925 Mus muscu
32	91	15.1	626	1	AL681532	AL681532 AL681532
33	91	15.1	787	8	DR849993	DR849993 JGI_CABE1
34	91	15.1	797	8	CX940605	CX940605 JGI_CAO7
35	91	15.1	811	3	B1526650	B1526650 602925809
36	91	15.1	847	8	DR873626	DR873626 JGI_CABH6
37	91	15.1	868	8	CX472102	CX472102 JGI_X2G53
38	91	15.1	910	8	CX849277	CX849277 JGI_CAL6
39	91	15.1	1505	10	AG429823	AG429823 Mus muscu
40	90.5	15.0	684	7	CO131269	CO131269 GR_EB43H
41	90.5	15.0	773	3	BU077405	BU077405 BU077405
42	90.5	15.0	816	7	CO113362	CO113362 GR_EB013
43	90.5	15.0	907	2	BG394138	BG394138 602456317
44	90.5	15.0	1151	6	CD496499	CD496499 CDA22-F06
45	90	14.9	487	3	BM249657	BM249657 K0836H12-

ALIGNMENTS

RESULT 1
T09652
LOCUS
DEFINITION
0224m3 gmbpFHB3.1, G. Roman Reddy Plasmodium falciparum genomic
clone 0224m, genomic survey sequence.
ACCESSION
T09652
VERSION
T09652.1 GI:319484
KEYWORDS
GSS.
SOURCE
Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM
Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 317)
Reddy,G.R., Chakrabarti,D., Schuster,S.M., Ferl,R.U., Almira,E.C.
and Dame,J.B.
Gene sequence tags from Plasmodium falciparum genomic DNA fragments
prepared by the genease activity of mung bean nuclease
Proc. Natl. Acad. Sci. U.S.A. 90, 9867-9871 (1993)
8234327
COMMENT
Other GSSs: 0224m7
Contact: Dame JB
Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.vetmed.ufl.edu
Seq primer: T3
Class: Shotgun.
Location/Qualifiers
1..317
/organism="Plasmodium falciparum"

FEATURES

source

/mol_type="genomic DNA"
/db_xref="taxon:5833"
/clone="0224m"
/lab_host="E. coli XLI-Blue"
/clone_lib="gmbpFHB3.1, G. Roman Reddy"
/note="Vector: pBluescript SK(+); Genomic DNA, from
asynchronous blood stage parasites of the cloned Honduran
HB3 isolate cultured in vitro, was digested with mung bean
nuclease in the presence of 30% formamide at 50°C
(Verneik, K.D., Imbereki, R.B., and McCutchan, T.F. 1998.
Nucleic Acids Research 16:6883-6896). The ends of the
fragments were ligated using T4 DNA polymerase, and the
dephosphorylated pBluescript SK(+). Recombinant plasmids
transformed E. coli XLI-Blue."

ORIGIN

Alignment Scores:
Pred. No.: 148e-36 Length: 317
Score: 396.00 Matches: 77
Percent Similarity: 96.25% Conservative: 0
Best Local Similarity: 96.25% Mismatches: 3
Query Match: 65.67% Indels: 0
DB: 10 Gaps: 0

US-09-837-344-20 (1-117) x T09652 (1-317)

QY 38 GtluileproAlaIleGluLeuProSerGluAenGluArgLyTyrtleProHisGln 57
|||
DB 1 GAATCGAAGCTATAGAACTTCATCAGAGATGAACGTGATATATATACCATCA 60
QY 58 SerSerLeuProGluInAerAenAerGlyAenSerAerAerSerLyGluIleSerIle 77
61 TCTTCTTACCTCAGACAGAACAGAGGATAGAGATTCCAGGAAATATCTATAA 120
QY 78 GtluVtHAsnAerGluSerIleThrTrpAenValGluYArgAerAerIleHisLy 97
121 GAAAAAACAATAGAAATCTTATTACAAATATTGTAAGAGACGAAAGATATACATA 180
DB 181 GGACATCTTGAAGAAAGAAAGATGTTCAATTAACCAAGTACAAAGAAATATCT 240
QY 98 GtluHleuGluGluLyAerGlySerIleLyProGluGluLyGluAerLySer 117
|||
DB 181 GGACATCTTGAAGAAAGAAAGATGTTCAATTAACCAAGTACAAAGAAATATCT 240

RESULT 2
BF133651 1782 bp mRNA linear EST 24-OCT-2000
LOCUS 601778550P1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4006597 5',
DEFINITION mRNA sequence.
ACCESSION BF133651
VERSION BF133651.1 GI:10972691
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1782)
NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM9238 row: c column: 14
High quality sequence stop: 550.
Location/Qualifiers
1..1782

FEATURES
SOURCE

/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:4006597"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu30"
/note="Organ: lung; Vector: pcMV-SPO6; Site 1: Nci1;
Site 2: Sal1; transgenic model MN1-1, expression driven by
Mtv-Ltr enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Alignment Scores:
Pred. No.: 0.319 Length: 1782
Score: 103.50 Matches: 31
Percent Similarity: 46.09% Conservative: 22
Best Local Similarity: 26.96% Mismatches: 39
Query Match: 17.16% Indels: 23
DB: 2 Gaps: 3

US-09-837-344-20 (1-117) x BF133651 (1-1782)

QY 2 GtnglugInglInAerLeuGluGluAerLyAerLyAerLyAerLyAerLyAerLy 21
|||
DB 1194 CAAGAGCAGACAGCAGAGAGGAGACACAAAGAAAGAAAGAAAGAAAGAAAG 1253
QY 22 LyAerLyGluHleuAerPleuAerIleuAerLeuAerLyGluAerLyGluAer 41
1254 AGACAAAGAAACAGACAGACAGAAACAAAGAGGACAAACAAAGAGATAGACCAAGAGCA 1313
QY 42 IleGluLeuProSerGluAenGluArgLyTyrtleProHisGlnSerSerLeuPro 61
1314 GCGCAACAGCAACGAAACGAAAGCA----- 1343
DB 62 GtlnAerAerAerGluAerSerAerAerSerLyGluIleSerIleGluLySerAen 81
1344 ---GAGAGAGAAAGAAAGAAAGAG-----AAGAGAGAGC 1379
QY 82 ArgGluSerIleThrAenValGluYArgAerAerIleHisLyGluHleuGlu 101
DB 1380 AGAGAAAGAAACAGAGCGCAAGAAAGAAAGAAAG-----AAGAGAGAG 1424
QY 102 GtluLyAerAerGlySerIleLyProGluGluLyGluAerLy 116
DB 1425 GAGAGCAAGAAAGAAAGTACAGACAGAGAGAAAGAAAGAACAG 1469

RESULT 3
AG332417/c 2131 bp DNA linear GSS 18-DEC-2004
LOCUS Mus musculus molossinus DNA, clone:MSHg01-122J15.T7, genomic survey
DEFINITION sequence.
ACCESSION AG332417
VERSION AG332417.1 GI:47905727
KEYWORDS GSS.
SOURCE Mus musculus molossinus (Japanese wild mouse)
ORGANISM Mus musculus molossinus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2131)
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
2 (bases 1 to 2131)
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.

TITLE	Direct Submission
JOURNAL	Submitted (17-NOV-2003) Meeahira Hattori The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:hattori@gsc.riken.jp, URL:http://nsp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT	Clones are derived from the mouse BAC library MSGM01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Teikoku Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp
PRIMERS	
Sequencing :	T7
LIBRARY	: pBACE3.6
Vector	: EcoRI
R.Site 1	: EcoRI
R.Site 2	:
FEATURES	Location/Qualifiers
Source	1..2131
	/organism="Mus musculus molossinus"
	/mol_type="genomic DNA"
	/sub_species="mossinus"
	/db_xref="taxon:57486"
	/clone="MSGM01-122J15.T7"
	/sex="male"
	/tissue_type="mixture of kidney and spleen"
	/clone_lib="MSGM01 Mouse Male BAC Library"
ORIGIN	
Alignment Scores:	
Pred. No.:	1.18 Length: 2131
Score:	99.50 Matches: 44
Best Similarity:	45.86% Conservative: 17
Percent Similarity:	33.08% Mismatches: 41
Query Match:	16.50% Indels: 32
DB:	Gaps: 7
US-09-837-344-20 (1-117) x AG332417 (1-2131)	
Oy	2 GlnGInGInGInArGArApRLeuGlunIarPlySaIAApThrLyLeAseLnuGluuarg 21
Dd	1192 CAAGAAGAGCAGCACACATCATCGAAAGAAAGAAAGAACCCACAGAAAGAAAGCAAGA 1133
Oy	22 LySLySGlu-----HisGLYAspILeuLeu 29
Dd	1132 GAABAAAGAAAGATCAAGAAAGACAGAAACGAATAATTAACGACAAAGGCACAACGAGA 1073
Oy	30 AlAgIuaRpLeuTyrgLYArGLEuGluIleProAlalIEglueuProSerGluabnglu 49
Dd	1072 GCACAAAAGATGGAGACAGACAGACAGATTACGAGCAA-GAGAGAAAGACGAGAGCGAGA 1014
Oy	50 ArGgLYrTYrTlleProHlbgInsereSeuProGlnaarPaenArhGlyYAmserArg 69
Dd	1013 AGA-----CAAAcCGAcGAGAGAGAAcCGCGAG 984
Oy	70 -----AapSerLyEGluIleSerIleIlegluLySTTrAmArGgluSerIleThr 86
Dd	993 AAAAAAATTCAGAGAAAGAGACATGACAAAGAAAGAGAACCAAGAA-----ACG 930
Oy	87 ThraenValGluGlyArGArGaRPleHllySLyGlyNIvleu-----GluGluLyelys 104
Dd	929 ACACACACTTAAGAACAGAAAGACAGACAAAGAA---CACACAAAAAGAGAGAGAAAG 873
Oy	105 AapGly----SerIleLyPProGluGluIuLySGluAaplys 116
Dd	872 GACAGCATTAAGCAAAAAAGAGACAAAAAACAACGAG 834
RESULT 4	
LOCUS	Cv271474/C 589 bp mRNA linear EST 22-SEP-2004
DEFINITION	WS0154.B21_M21 PTXN-IB-A-6 Populus trichocarpa x Populus nigra cDNA

```

ACCESSION   Clone WS0154_M21 3', mRNA sequence.
VERSION     CV271474
KEYWORDS    CV271474.1 GI:52524449
SOURCE      EST.
ORGANISM    Populus trichocarpa x Populus nigra
            Populus trichocarpa x Populus nigra
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1 (bases 1 to 589)
AUTHORS     Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
            Kirkpatrick,R., Liu,J., Palmquist,D., Stott,U., Barber,S., Yang,G.,
            Babakoff,R., Brown-John,M., Chand,S., Featherstone,R., Maason,A.,
            Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
            Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
            Bohlmann,J.
TITLE       The poplar transcriptome: Analysis of expressed sequence tags from
            multiple cDNA libraries
JOURNAL     Unpublished (2004)
COMMENT     Contact: Joerg Bohlmann
            Genome BC forest genomics program
            University of British Columbia
            Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
            Vancouver, British Columbia, Canada. V6T 1Z3
            Tel: 1-604-822-0282
            Fax: 1-604-822-2114
            Email: bohlmann@msl.ubc.ca
            Plate: WS0154 row: M column: 21
            High quality sequence stop: 589
            POLYA=Yes.

FEATURES             location/Qualifiers
     source           1..589
                     /organism="Populus trichocarpa x Populus nigra"
                     /mol_type="mRNA"
                     /culivar="NXM6"
                     /db_xref="taxon:293756"
                     /clone="WS0154_M21"
                     /sex="male"
                     /lab_host="E. coli DH10B T1 phage resistant cells"
                     /clone_id="PrXN-IB-A-6"
                     /note="vector: plusscript II SK (+) XR; Site_1: EcoRI (5'
end of cDNA); Site_2: XhoI (3' end of cDNA); Sapling trees
two metres in height and grown under greenhouse conditions
were exposed to continuous feeding by Cryophrynus
lapathi (poplar and willow borer) adults caged on the
sapling using mesh bags. Bark (with phloem and cambium
attached) from within the caged region was harvested 2
hours, 6 hours and 48 hours after the onset of treatment.
mRNA was isolated from each tissue source independently
and equal quantities of mRNA from each tissue were then
pooled. cDNA was prepared from 5 micrograms of mRNA and
directionally ligated into the pluscript II SK (+) XR
vector using the pluscript II XR cDNA library
Construction Kit according to manufacturer's instructions
with modifications (Stratagene). Plasmid DNA was then
transformed by electroporation into DH10B cells
(invitrogen) for propagation."

ORIGIN
Alignment Scores:
Pred. No.:          0.356              Length:          589
Score:              98.00              Matches:         34
Percent Similarity: 47.90%              Conservative:    23
Best Local Similarity: 28.57%            Mismatches:     52
Query Match:        16.25%              Indels:         10
DB:                  7                  Gaps:           4

US-09-837-344-20 (1-117) x CV271474 (1-589)
QY      2 GInGluGInGlnAArgPLeuGlUGlnGlnAArgLySLysAlaApThrLYtLysAsnLeuGUArg 21
       ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      518 AAAGGAAGAAGAACTGAAGCAAGAAAAGCAAAGAACGACGACAAAAAGAAAT---GAAACA 462

```

```

QY      22  Lyslysglyhislglaspriileuallagluapleuylrglyargyleuqluileproala 41
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      461  GATGCTGMAAATGTTGAT-----GCAAGTATGTTTGTGGCCACAGATGACAAAAG 408
QY      42  lileguleuProsergluasngluarglylrylrylreolileglnserleu--- 60
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      407  AGGMA-----AAGCACAAGATAGAAACACCGAAGACGATCAAGTGCCTTAT 354
QY      61  -----ProglinaPasnarglyasnserlryasnserlrygluileserile 76
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      353  GATGTAATCTGCAGAGATGAGAAAGAGTCTAAATAATCGCAGCATGACGCT 294
QY      77  lilegulythrAsnarglyserilerThrAsnValgluGlyArgArgAspillehis 96
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      293  GACCTTAAATAATCTGAAAGACACATACACCTGATCATGATGCGAAAGCAGCAT 234
QY      97  LysglYhislleugluqluLyslyAspqlYserilerleProgluGlnlysgluasp 115
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      223  AAAAGCACAAGAGATCACCGTGTGTTCCGCAAAATGCGAGTAAATGAGAG 177

RESULT 5
CA082727 605 bp mRNA linear EST 23-SEP-2003
LOCUS SCBFAM2025D06.g AM2 Saccharum officinarum cDNA clone SCBFAM2025D06
DEFINITION 5', mRNA sequence.
ACCESSION CA082727
VERSION CA082727.1 GI:34936038
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE 1 (bases 1 to 605)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLES The libraries that made SUCSEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: paruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 025 row: D column: 06
Seq primer: 77 Promoter Primer.
FEATURES
    source
        1..605
            /organism="Saccharum officinarum"
            /mol_type="mRNA"
            /db_xref="taxon:4547"
            /clone="SCBFAM2025D06"
            /lab_host="DH10B"
            /clone_lib="AM2"
            /note="Organ: Apical meristem and tissues surrounding of
            immature plant; Vector: pSPORT1, Site 1: SalI, Site 2:
            NotI; An unidirectional cDNA library generated from
            [Apical meristem and tissues surrounding of immature
            plant]. cDNA was prepared from polyA+ mRNA using
            Superscript Plasmid System Kit (Invitrogen). The
            double-strand cDNAs were fractionated in a sepharose
            CL-2B 40cm-columns and fragments sizing between 0.8 and
            1.5 Kb were directionally cloned into the vector. Details
            of each source of RNA and library construction can be
            obtained at http://sucsest.lad.ic.unicamp.br/public"

```

```

Score: 97.50
Percent Simlarity: 49.59%
Best Local Similarity: 26.45%
Query Match: 16.17%
DB: 6 Gaps: 5

US-09-837-344-20 (1-117) x CA082727 (1-605)

QY      2  GlnGlnGlnGlnAArgApLeuGlnGlnArgYsAlaSpThrYsAsnLeuGlnArg 21
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      76  GAAAAACAAGAGAGGGGCAAGACAAAGAAAAAGCAAGACGAGAAAGATGGAACAG 135
QY      22  Lyslysgluhis-----GlyAspIleuallagluapleuylrglyarg 36
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      136  AAGAGGGCCATCTCGAATATGAAATGTAAGTGAACCCGACAGAGATGTTCCATGAA 195
QY      37  leuqluileproalaileguleuProsergluasngluarglylrylrylreohis 56
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      196  AAGAGTTGGCAGAGAGCTGCAGACTCGGGACCCGTGA-----ATCAAGCTA 243
QY      57  GlnserleuProglinaPasnarglyasnserlryasnserlrygluileserile 76
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      244  ACTGACTGAACCAACAGAGAAAGAAATGATGTGAAGTTGCAAGGC----- 294
QY      77  lilegulythrAsnarglyserilerThrAsnValgluGlyArgArgAspille 95
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      295  ---AAGAAAGAGAACAAAGAGAGGACCAAGGATGATCAAGAAAGAAAGAGACGTT 351
QY      96  HlslysglyhislleugluqluLyslyAspqlYserilerleProgluGlnlysgluasp 115
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      352  GGTACGGCGCATGAGAGGTGAGAAAGAAACAAAGAA--AAGAAAGAGAGAGAGAGAC 408
QY      116  Lys 116
Db      409  AAG 411

RESULT 6
DR983403 684 bp mRNA linear EST 03-AUG-2005
LOCUS JGI_AOKF1090.rev AOKF Acropora palmata spawned eggs Acropora
DEFINITION JGI_AOKF1090.rev AOKF Acropora palmata spawned eggs Acropora
ACCESSION DR983403
VERSION DR983403.1 GI:71778109
KEYWORDS EST.
SOURCE Acropora palmata
ORGANISM Acropora palmata
Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Scleractinia;
Astrocoenina; Acroporidae; Acropora.
REFERENCE 1 (bases 1 to 684)
AUTHORS Schwarz,J.A., Brokslein,P., Manohar,C., Coffroth,M.A., Szmant,A.
and Medina,M.
TITLES Coral-Symbiodinium EST Project
JOURNAL Unpublished (2005)
COMMENT Other ESTs: JGI_AOKF1090.fwd
Contact: Schwarz, JA, Medina, M.
Evolutionary Genomics
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925-296-5823
Email: jaszczewski@lbl.gov
cDNA Library Preparation: DOE Joint Genome Institute:
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone Id and the direction of sequencing. The suffix '.rev'
indicates a reverse sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Poly-T: A run of 14 or more T residues at the beginning of this
sequence has been removed.
Plate: AOKF 0009 row: d column: 10
High quality sequence stop: 546.
FEATURES
    source
        1..684
            Location/Qualifiers

```

/organism="Acropora palmata"
/mol_type="mRNA"
/db_xref="taxon:6131"
/clone="AOKF1090"
/dev_stage="Freshly spawned eggs"
/lab_host="Electromax DH108"
/note="Vector: pMDR-LIB, Site 1: SfiI, Site 2: SfiI, The library was prepared from total RNA using the Creator SMART cDNA Library Construction Kit with the LD-PCR method to amplify the cDNA. Amplified cDNA was digested with SfiI, size selected for >400bp, and ligated into the pMDR-LIB vector."

ORIGIN

Alignment Scores:

Pred. No.:	0.737	Length:	684
Score:	96.00	Matches:	40
Percent Similarity:	43.14%	Conservative:	26
Best Local Similarity:	26.14%	Mismatches:	33
Query Match:	15.92%	Indels:	54
DB:	8	Gaps:	8

US-09-837-344-20 (1-117) x DR983403 (1-684)

QY 1 LeuGlnGlnGlnGlnAArgPLeuGlnAArgLysAlaAsp----- 14
DB 486 TTGCMAAGCAGCAGTAACTCGAAAGAGGCTGAGCTTGAAGTCCACTCT 427
QY 15 -----ThrlYsLysAanLeuGlu-----ArgLysLysGluHtIsGlyAsp----- 27
DB 426 TACTTCAGAGAAAGCTTCTGTAATTCAGAACAGACAGAACAGATGGCGAGAAAGA 367
QY 28 ---IleLeuAlaGluAspLeuTyrgLyArgLeuGluIleProAlaIleGluLeuProSer 46
DB 366 AAGCTCACAGACAGCAGATTGTGACAGCAAACTGAATC-----CTACATCA 319
QY 47 GLuAanGluAArgGlyTyrrTyrlleProHtIsGlnSerLeuProGlnAanPAnArgGly 66
DB 318 GAACACGAA-----ACATTAATCCAGTTTG----- 292
QY 67 AAnSerArgAspSerLysGluIleSerIleIleGlu----- 78
DB 291 GGGAGTCAGATGCACTGATTTGCGTTCGGAAGAAAAGCAAGTTCCTGTGACA 232
QY 79 ---LysThrAanArgLysSerIleThrThraenValGluArgArgAsp----- 94
DB 231 ATTAAAGTCTCGAGATAAAGTGTGATCAAGCAAACTGTAAAGACCTACAGAG 172
QY 95 -----IleHtIsLysGluHtIsLeuGlu 101
DB 171 AAGTCGAAATGATGAACGAAATTAAGACTTTTCAGATCGAAGAAAGCCCTTGAA 112
QY 102 GLuLysLysAanArgLysSerIleLysProGlnGlnLysGlu 114
DB 111 AAGGAAAAGATGTTTCGTCTCGCGGAAAAGTCAGGAA 73

RESULT 7
B1559491 1042 bp mRNA linear EST 05-SEP-2001
LOCUS 60325811.F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:525338 5',
DEFINITION mRNA sequence.
ACCESSION B1559491
VERSION B1559491.1 GI:15446805
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1042)
REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@email.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshitaki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1A11747 row e column: 03
High quality sequence stop: 12.
Location/Qualifiers

FEATURES

Source

1..1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5295338"
/lab_host="DH108"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gagag); Oligo-dT primed using primer
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.2 kb and normalized to R0F 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci,
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.:	1.25	Length:	1042
Score:	96.00	Matches:	36
Percent Similarity:	41.48%	Conservative:	20
Best Local Similarity:	26.67%	Mismatches:	57
Query Match:	15.92%	Indels:	22
DB:	3	Gaps:	5

US-09-837-344-20 (1-117) x B1559491 (1-1042)

QY 3 GlnGlnGlnAArgPLeuGlnAArgLysAlaAspThrlYsLysAanLeuGluAArgLys 22
DB 474 CAAGAAAACACCAAGCAAAAGACGAAACGACCAAAAGCAAAACAGAGACAA 533
QY 23 LysGluHtIsGlyAsp---IleLeuAlaGluAspLeuTyrgLyArgLeuGluIleProAla 41
DB 534 AGGACGAAAGCGACAGATGAAGAAAGAGACGAGGAAAGGATGAAGAACGGGG 593
QY 42 IleGluLeuProSerGluAanGluAArgGlyTyrrTyrlleProHtIsGlnSerLeuPro 61
DB 594 AGGGA---CCAGACAAAGACGAGCGGGAAGCAAAACAGAAACAAAGAGA 650
QY 62 GlnAanPAnArg-----GlyAanSerArgAspSerLysGluIleSerIleIleGluLys 79
DB 651 CAGACAAAGAAAACACAGAAAGAAAGAGAAAACAAACACAAAAGACGAAAG 710
QY 80 ThrAanArgLysSerIleThrThraenValGluArgArgAspIleHtIsLys----- 97
DB 711 AGACACGAACACCGACGACGAGAAAGAAAGGAAAGCAAGAACACCAAGACCA 770
QY 98 -----GlyHtIs-----LeuGlu 101
DB 771 GGGAAACGAGACACAGAGAAAGACGACACGAGAGAAAGAGAGAAAGAAAGAA 830
QY 102 GLuLysLysAanArgLysSerIleLysProGlnGlnLysGluAanPys 116
DB 831 GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 875

RESULT 8
BG497835

Qy 91 G1yArGaRgArp1leH1eLyeG1yH1eLyeG1uG1uLySeArpG1ySeR1leLyePro 110
 Db 478 CAAAGATGACGATTTGAAAGAAAAGAAAAGAAAAGTTTGAATACTAAGAG 537
 Qy 111 G1uG1uLySeArp1le 116
 Db 538 GAACAAAACAATCTGAA 555
 RESULT 13
 CA234102 974 bp mRNA linear EST 25-SEP-2003
 LOCUS SCEZAM2250D06.9 AM2 Saccharum officinarum cDNA clone SCEZAM2250D06
 DEFINITION 5', mRNA sequence.
 ACCESSION CA234102
 VERSION CA234102.1 GI:35301309
 KEYWORDS EST.
 SOURCE Saccharum officinarum
 ORGANISM Saccharum officinarum
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.
 REFERENCE 1 (bases 1 to 974)
 AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
 TITLE The libraries that made SUCEST
 JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 COMMENT Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parvada@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bcccenter.fcav.unesp.br
 Plate: 250 row: D column: 06
 Seq primer: T7 Promoter Primer.
 Location/Qualifiers
 1..974
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCEZAM2250D06"
 /lab_host="DH10B"
 /clone_1lb="AM2"
 /note="Organ: Apical meristem and tissues surrounding of
 immature plants; Vector: pSport1; Site 1: SalI; Site 2:
 NotI; An unidirectional cDNA library generated from
 (Apical meristem and tissues surrounding of immature
 plants). cDNA was prepared from polyA+ mRNA using
 SuperScript Plasmid System Kit (Invitrogen). The
 double-strand cDNAs were fractionated in a sepharose
 CL-2B 40cm-columns and fragments siting between 0.8 and
 1.5 Kb were directionally cloned into the vector. Details
 of each source of RNA and library construction can be
 obtained at http://succest.fad.ic.unicamp.br/public"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.72 Length: 974
 Score: 94.50 Matches: 31
 Percent Similarity: 49.59% Conservative: 29
 Best Local Similarity: 25.62% Mismatches: 46
 Query Match: 15.67% Indels: 15
 Gaps: 5
 US-09-837-344-20 (1-117) x CA234102 (1-974)

Qy 2 G1nG1uG1nArp1leG1uG1uLySa1eArp1leYsArnLeuG1uArG 21
 Db 47 GAAAAACAGAAAGCAAGAAAGCAAGAAAGCAAGAGCAAGATGGAACAG 106

Qy 22 LyELySg1uH1e-----G1yArp1leuH1eLyeLag1uArpLeuTyRg1yArG 36
 Db 107 AAGAGGGCCATCTCGAGAAATTAAGATGCTGAAGCCGACAGAAAGATGTCATGAAC 166
 Qy 37 LeuG1u1ePro1a1leG1uLeuProSeRg1uAsnG1uArG1yTyR1leProH1e 56
 Db 167 AAAGAGTTGGACAGAGCTGCAGACTCGGGAGCCGTCGA-----ATCAAGCTA 214
 Qy 57 G1nSeRse1euproG1nArp1leArp1leYsArnSeR1eArp1leSeR1le 76
 Db 215 ACTGACACTTAACCCACGAGAGAAAGAAATTAATGTAAGATTCAGAAAGC----- 265
 Qy 77 1leG1uYrThraRnaArG1uSeR1leThrThraSna1---G1uG1yArGaRgArp1le 95
 Db 266 ---AAGAAAGAAACAAAGAGAGAGACAGACAGATGAAGTCAAGAAAGAAAGAGAGCTT 322
 Qy 96 H1eLySg1yH1eLyeG1uG1uLySa1eArp1leYsArnLeuG1uArG1yTyR1le 115
 Db 323 GGTACAGGCGCATGAAGTGAAGAAAGCAAGAA---AAGAAAGAGAAAGAGAGAGC 379
 Qy 116 Lys 116
 Db 380 AAG 382
 RESULT 14
 BG284329 1665 bp mRNA linear EST 21-FEB-2001
 LOCUS BG284329
 DEFINITION 602408426F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4520407 5',
 mRNA sequence.
 ACCESSION BG284329
 VERSION BG284329.1 GI:13035166
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE 1 (bases 1 to 1665)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.llnl.gov
 Plate: LHM10418 row: d column: 08
 High quality sequence stop: 170.
 Location/Qualifiers
 1..1665
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4520407"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_91"
 /note="Organ: prostate; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC library."

ORIGIN
 Alignment Scores:
 Pred. No.: 3.36 Length: 1665
 Score: 94.50 Matches: 34
 Percent Similarity: 40.00% Conservative: 18
 Best Local Similarity: 26.15% Mismatches: 45

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:28:45 ; Search time 2602.64 Seconds
(without alignment)
2555.359 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LDEQQRDEQRKADTKKNLE.....GHLEKKQGSIRPEQKEDKS 117

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xjh
-O/cgcn2.1/USPTO.spool/h/US09837344/runat.05122005.133532.734/app.query.fasta.1.718
-DB=GenEmbl -QPM=fastcap -SUFFIX=age -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdt -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=prc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09837344 @CGCN 1.1 5142 @runat.05122005.133532.734 -NCPUs=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_pro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vl.*
14: gb_hgt.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	603	100.0	950	A28740 CDNA for LS
2	603	100.0	950	ARI62961 Sequence
3	603	100.0	950	ARI78250 Sequence

4	603	100.0	1482	6	ARI62970 Sequence
5	603	100.0	1482	6	ARI78259 Sequence
6	603	100.0	1493	6	A28743 CDNA for LS
7	603	100.0	1493	6	ARI62967 Sequence
8	603	100.0	1493	6	ARI78256 Sequence
9	603	100.0	1496	6	PFLSA13
10	602	99.8	1368	11	AF751501
11	602	99.8	5970	2	PFLSA1G
12	600	99.5	1482	6	ARI62974
13	600	99.5	1482	6	ARI78263
14	597	99.0	253001	2	AE014834
15	554	91.9	843	2	PFLSAC
16	554	91.9	843	2	PFLSAH
17	554	91.9	843	2	PFLSAI
18	553	91.7	843	2	PFLSAB
19	553	91.7	843	2	PFLSAB
20	553	91.7	843	2	PFLSAF
21	553	91.7	843	2	PFLSAK
22	550	91.2	843	2	PFLSAG
23	550	91.2	843	2	PFLSAL
24	549	91.0	843	2	PFLSAIA
25	549	91.0	843	2	PFLSAIB
26	549	91.0	843	2	PFLSAIC
27	549	91.0	843	2	PFLSAIN
28	549	91.0	843	2	PFLSAD
29	549	91.0	843	2	PFLSAJ
30	549	91.0	843	2	PFLSAM
31	549	91.0	843	2	PFLSAN
32	549	91.0	843	2	PFLSAP
33	469	77.8	795	2	AF246996
34	224	37.1	126	2	AY090552
35	104.5	17.3	161790	14	AC022371
36	104.5	17.3	161863	8	AL606970
37	103.5	17.2	162642	5	BX908804
38	102	16.9	13097	1	AE000673
39	102	16.9	218208	14	AC073700
40	101.5	16.8	110000	15	CR382133_11
41	98.5	16.3	107112	8	AL596210
42	98.5	16.3	184315	14	AC073482
43	98.5	16.3	347050	2	PFA929351
44	97	16.1	2130	6	AR549811
45	97	16.1	112968	8	AC093833

ALIGNMENTS

RESULT 1	A28740	950 bp	DNA	linear	PAT 04-JUN-1995
LOCUS	A28740				
DEFINITION	CDNA for LSA-R-NR protein.				
ACCESSION	A28740				
VERSION	A28740.1	GI:1247512			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 950)				
JOURNAL	LIVER-STAGE-SPECIFIC PEPTIDE SEQUENCES OF P. FALCIPARUM BEARING				
FEATURES	EPITOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES				
source	Patent: WO 921384-A 29 20-AUG-1992;				
	Location/Qualifiers				
	1..950				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN
Alignment Scores: 1,31e-50 Length: 950
Pred. No.: 603.00 Matches: 117
Score:

VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	ORIGIN
AR162970.1	GI:16233435	Unknown. Unclassified.	1 (bases 1 to 1482)	Guerin-Marchand,C. and Druille,P.	Peptide sequences specific for the hepatic stages of <i>P. falciparum</i> bearing epitopes capable of stimulating the T lymphocytes	Patent: US 6270771-A 41 07-AUS-2001; Location/Qualifiers	1..1482 /organism="unknown" /mol_type="unassigned DNA"	
Alignment Scores:			2.09e-50	Length: 1482				
Pred. No.:			603.00	Matches: 117				
Score:			100.00%	Conservative: 0				
Percent Similarity:			100.00%	Mismatches: 0				
Best Local Similarity:			100.00%	Indels: 0				
Query Match:			100.00%	Gaps: 0				
DB:			6					
US-09-837-344-20 (1-117) x AR162970 (1-1482)								
QY	1	LeuGInGInGInGInaRgApLeuGInGInaRgLyLaApRThrLyLyAsnLeuGlu 20						
Db	610	TTGCAGAGAGCAGCAAGATTAGAACAAAGAAAGCGCTGATACGAAAAAATTTAGAA 669						
QY	21	ArgLyLyLyGluHisGlyAspRLeLeuLAsGlyAspLeuTyRgLyArgLeuGluLeuPro 40						
Db	670	AGAAAAAGAAACATGAGATATATTACAGAGGATTTATATGGCTGTGAAATACCA 729						
QY	41	AlaIleGluLeuProSerGluAsnGluArgGlyTyRyIleProHisGlnSerSerLeu 60						
Db	730	GCTATAGAACTTCCTCATCGAAATATAGCTGATTTATATACCAATCATCTTCCTTA 789						
QY	61	ProGlnaPAsnArgLyAsnSerArgAspSerLyGluIleSerIleIleGluLyThr 80						
Db	790	CCTCAGCAACAAGAGGGAATGTAAGATTCGAAAGAAATATCTATATGAAAAACA 849						
QY	81	AsnArgGluSerIleThrThraAsnValGluGlyArgGAspRLeHisLyGlyHisLeu 100						
Db	850	AATAGAGAACTCATTTACAACAATGTTGAAAGACCAAGGATATCATTAAGACATCTT 909						
QY	101	GluGluLyLyLyAspGlySerIleLyProGluGlnLyGluAspLySer 117						
Db	910	GAAAGAAAGAAAGATGTTCAATTAAACCAAGAACAAAGAAATTAATCTT 960						
RESULT 5								
AR178259			1482 bp	DNA	linear	PAT 20-APR-2002		
LOCUS	AR178259							
DEFINITION	Sequence 41 from patent US 6319502.							
ACCESSION	AR178259							
VERSION	AR178259.1							
KEYWORDS	GI:20219397							
SOURCE								
ORGANISM	Unknown.							
REFERENCE	Unclassified.							
AUTHORS	1 (bases 1 to 1482)							
TITLE	Guerin-Marchand,C. and Druille,P.							
JOURNAL	Peptide sequences specific for the hepatic stages of <i>P. falciparum</i>							
FEATURES	bearing epitopes capable of stimulating the T lymphocytes							
source	Patent: US 6319502-A 41 20-NOV-2001; Location/Qualifiers							
ORIGIN	1..1482 /organism="unknown" /mol_type="unassigned DNA"							
Alignment Scores:			2.09e-50	Length: 1482				
Pred. No.:			603.00	Matches: 117				
Score:			100.00%	Conservative: 0				
Percent Similarity:			100.00%	Mismatches: 0				
Best Local Similarity:			100.00%	Indels: 0				
Query Match:			100.00%	Gaps: 0				
DB:			6					
US-09-837-344-20 (1-117) x AR162970 (1-1482)								
QY	1	LeuGInGInGInGInaRgApLeuGInGInaRgLyLaApRThrLyLyAsnLeuGlu 20						
Db	610	TTGCAGAGAGCAGCAAGATTAGAACAAAGAAAGCGCTGATACGAAAAAATTTAGAA 669						
QY	21	ArgLyLyLyGluHisGlyAspRLeLeuLAsGlyAspLeuTyRgLyArgLeuGluLeuPro 40						
Db	670	AGAAAAAGAAACATGAGATATATTACAGAGGATTTATATGGCTGTGAAATACCA 729						
QY	41	AlaIleGluLeuProSerGluAsnGluArgGlyTyRyIleProHisGlnSerSerLeu 60						
Db	730	GCTATAGAACTTCCTCATCGAAATATAGCTGATTTATATACCAATCATCTTCCTTA 789						
QY	61	ProGlnaPAsnArgLyAsnSerArgAspSerLyGluIleSerIleIleGluLyThr 80						
Db	790	CCTCAGCAACAAGAGGGAATGTAAGATTCGAAAGAAATATCTATATGAAAAACA 849						
QY	81							

Score:	603.00	Matches:	117
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-837-344-20 (1-117) x AR178259 (1-1482)			
QY	1 LeuGInGluGInGInAArgApLeuGluGInAArgLyAa1AaPThrLyAaALeuGlu 20		
DB	610 TTGCAGAGCGACGCAAAAGATTATAGAACAAAGAGGCTGATACGAAAAAATTTCGA 669		
QY	21 ArgLyAaLySGluNH1SGLyAaPLeuA1aGluAaPLeuTYrGlyAArgLeuGlu1LePro 40		
DB	670 AGAAAAAAGGAACATGGAGATATATTTCGACAGAGATTATATAGTCTGTTTAAATACCA 729		
QY	41 A1a1leGluAeuProSerGluAaNGluAArgLyTYrTYr1leProH1sGInSerSerLeu 60		
DB	730 GCTATAGAACTTCATCAGAAAAAGAACGGTGAATATATATACCATCATCTCTTTA 789		
QY	61 ProGInAaPaAaArgGlyAaAaSerAArgAaPSeSerLySGlu1leSer1le1leGluLyAThr 80		
DB	790 CCTCAGGCAACAAGAGGAAATAGTAGGATTCCAGAGAAATATCTATATATAGAAAAACA 849		
QY	81 AaAaArgGluSer1leThrAaAaValGluGlyAArgAaPLeuH1eH1eLySGluNH1eLeu 100		
DB	850 AATAGAGATATTATTACCAAAAGTTGAAAGGACAAAGGATATACATTAAGACATCTT 909		
QY	101 GluGluLyAaAaPGLySer1leLyAaProGluGluNHysGluAaPLeuLySer 117		
DB	910 GAAGAAAAGAAAGATGCTTCATTAATAACCAAAAGAAAGATTAATCT 960		
RESULT 6			
A28743	A28743	1493 bp	DNA linear PAT 04-JUN-1995
LOCUS	DEFINITION	CDNA for LSA gene 3' (partial).	
ACCESSION	A28743		
VERSION	A28743.1	GI:1247516	
KEYWORDS			
SOURCE			
ORGANISM			
	Homo sapiens (human)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo.		
	1 (bases 1 to 1493)		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
	1..1493		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.1e-50	Length:	1493
Score:	603.00	Matches:	117
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-837-344-20 (1-117) x A28743 (1-1493)			
QY	1 LeuGInGluGInGInAArgApLeuGluGInAArgLyAa1AaPThrLyAaALeuGlu 20		
DB	607 TTGCAGAGCGACGCAAAAGATTATAGAACAAAGAGGCTGATACGAAAAAATTTCGA 666		
QY	21 ArgLyAaLySGluNH1eGlyAaPLeuA1aGluAaPLeuTYrGlyAArgLeuGlu1LePro 40		
DB	667 AGAAAAAAGGAACATGGAGATATATTTCGACAGAGATTATATAGTCTGTTTAAATACCA 726		

QY	41	Ala11	egluLeuProse	gLUa	ngLUa	rgLUa	gYTrYTrIle	Prohi	egIn	Se	serLeu	60
Db	727	GCTAT	GAAC	TCTCCAT	CA	GAATA	TGAAC	GTGAT	TTTAT	ATAC	CAATCA	ATCTTCTTTA 786
OY	61	ProG	lna	Pe	a	Asn	ArgI	YAsn	Ser	r	gAsp	Ser
Db	787	CCTCA	GGAC	CA	CA	CGAG	GGAT	TTGT	GA	ATTC	CA	AGGA
OY	81	Asn	ArgI	Se	rIle	Thr	Asn	Val	gLU	gIa	gAr	gAsp
Db	847	AATAG	AGAT	CTTAT	TAC	AA	CAAT	GT	TGA	AGCA	GAAG	GAATAT
OY	101	Gluc	lu	Le	Val	Asp	gI	Le	Val	Pro	gI	Val
Db	907	GAG	GA	AA	GA	AA	GA	AT	GT	TTA	AA	CA
OY	117	Gluc	lu	Le	Val	Asp	gI	Le	Val	Pro	gI	Val
Db	957	GAG	GA	AA	GA	AA	GA	AT	GT	TTA	AA	CA

RESULT 7					
AR162967					
LOCUS	AR162967	1493 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION	Sequence 38 from patent US 6270771.				
ACCESSION	AR162967				
VERSION	AR162967.1	GI:16233432			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1493)				
TITLE	Guerin-Marchand, C. and Druilhe, P.				
JOURNAL	Peptide sequences specific for the hapatic stages of P. falciparum				
FEATURES	bearing epitopes capable of stimulating the T lymphocytes				
source	Patent: US 6270771-A 38 07-AUG-2001;				
	Location/Qualifiers				
	1..1493				

ORIGIN	/mol_type="unassigned DNA"
Alignment Scores:	
Pred. No.:	2.1e-50
Score:	603.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
	Gaps: 0
US-09-837-344-20 (1-117) x ARI62967 (1-1493)	

OY		1	L <u>e</u> uG <u>i</u> n <u>G</u> I <u>u</u> G <u>i</u> n <u>G</u> I <u>A</u> r <u>G</u> a <u>S</u> p <u>Le</u> u <u>G</u> i <u>n</u> a <u>T</u> g <u>Ly</u> sA <u>I</u> a <u>S</u> P <u>Th</u> r <u>Ly</u> e <u>S</u> y <u>a</u> n <u>Le</u> u <u>G</u> i <u>U</u>	20
D _B		607	T <u>T</u> G <u>C</u> A <u>A</u> G <u>A</u> G <u>C</u> A <u>C</u> A <u>A</u> G <u>A</u> G <u>A</u> T <u>T</u> T <u>A</u> A <u>C</u> A <u>A</u> G <u>A</u> G <u>C</u> G <u>T</u> G <u>A</u> T <u>A</u> C <u>G</u> A <u>A</u> A <u>A</u> A <u>A</u> A <u>A</u> T <u>T</u> T <u>A</u> G <u>A</u>	666
OY		21	A <u>R</u> g <u>L</u> y <u>V</u> l <u>E</u> g <u>I</u> N <u>H</u> I <u>E</u> G <u>I</u> A <u>S</u> p <u>I</u> l <u>E</u> u <u>A</u> g <u>I</u> a <u>S</u> p <u>Le</u> u <u>T</u> y <u>G</u> I <u>A</u> r <u>G</u> l <u>e</u> u <u>G</u> i <u>U</u> l <u>E</u> Pro	40
D _B		667	A <u>G</u> A <u>A</u> A <u>A</u> A <u>A</u> G <u>A</u> A <u>C</u> A <u>T</u> G <u>A</u> G <u>A</u> T <u>A</u> T <u>T</u> A <u>G</u> A <u>G</u> A <u>G</u> A <u>T</u> T <u>T</u> A <u>T</u> A <u>G</u> C <u>G</u> T <u>T</u> A <u>G</u> A <u>A</u> A <u>T</u> A <u>C</u> A	726
OY		41	A <u>A</u> l <u>I</u> e <u>G</u> I <u>U</u> l <u>E</u> u <u>P</u> ro <u>S</u> e <u>R</u> G <u>I</u> a <u>S</u> n <u>G</u> I <u>A</u> u <u>G</u> I <u>Y</u> T <u>T</u> Y <u>T</u> l <u>E</u> Pro <u>H</u> I <u>E</u> G <u>I</u> S <u>e</u> r <u>S</u> e <u>R</u> le <u>U</u>	60
D _B		727	G <u>C</u> T <u>A</u> T <u>A</u> G <u>A</u> A <u>C</u> T <u>T</u> C <u>A</u> T <u>C</u> A <u>G</u> A <u>A</u> A <u>T</u> A <u>A</u> C <u>T</u> G <u>A</u> T <u>T</u> A <u>T</u> A <u>T</u> A <u>C</u> A <u>C</u> A <u>T</u> C <u>A</u> A <u>C</u> T <u>T</u> C <u>T</u> T <u>A</u>	786
OY		61	P <u>R</u> O <u>G</u> I <u>A</u> S <p>Proteinase A</p> <u>P</u> ar <u>A</u> m <u>I</u> Y <u>A</u> S <u>n</u> S <u>e</u> r <u>I</u> r <u>G</u> a <u>S</u> p <u>S</u> e <u>R</u> l <u>E</u> g <u>I</u> U <u>S</u> e <u>R</u> I <u>E</u> l <u>E</u> g <u>I</u> U <u>S</u> Y <u>Th</u> r	80
D _B		787	C <u>C</u> T <u>C</u> G <u>A</u> G <u>A</u> C <u>A</u> C <u>A</u> G <u>A</u> G <u>A</u> T <u>G</u> T <u>A</u> G <u>A</u> K <u>A</u> T <u>T</u> C <u>C</u> A <u>A</u> G <u>A</u> A <u>T</u> T <u>T</u> C <u>T</u> A <u>T</u> A <u>T</u> A <u>G</u> A <u>A</u> A <u>A</u> A <u>C</u> A	846
OY		81	A <u>S</u> n <u>A</u> r <u>G</u> I <u>S</u> e <u>R</u> I <u>E</u> Th <u>r</u> Th <u>r</u> An <u>A</u> l <u>G</u> I <u>A</u> r <u>G</u> a <u>S</u> p <u>I</u> h <u>I</u> S <u>Ly</u> S <u>G</u> I <u>Y</u> h <u>I</u> S <u>Le</u> U	100
D _B		847	A <u>A</u> T <u>A</u> G <u>A</u> G <u>A</u> T <u>C</u> A <u>T</u> T <u>A</u> C <u>A</u> C <u>A</u> A <u>T</u> G <u>T</u> G <u>A</u> A <u>G</u> A <u>C</u> A <u>A</u> G <u>A</u> G <u>A</u> T <u>A</u> T <u>A</u> C <u>A</u> T <u>A</u> A <u>G</u> A <u>C</u> A <u>T</u> C <u>T</u>	906
OY		101	G <u>I</u> U <u>G</u> I <u>U</u> l <u>V</u> l <u>E</u> A <u>P</u> G <u>I</u> Y <u>S</u> e <u>R</u> I <u>E</u> l <u>Y</u> S <p>Proteinase B</p> <u>P</u> ro <u>G</u> I <u>U</u> G <u>I</u> N <u>U</u> l <u>V</u> G <u>I</u> A <u>S</u> p <u>Y</u> S <u>e</u> R	117
D _B		907	G <u>A</u> G <u>A</u> A <u>A</u> A <u>A</u> G <u>A</u> A <u>G</u> T <u>G</u> T <u>T</u> C <u>A</u> T <u>A</u> T <u>A</u> A <u>C</u> C <u>A</u> G <u>A</u> C <u>A</u> A <u>A</u> A <u>A</u> G <u>A</u> A <u>G</u> A <u>T</u> A <u>A</u> A <u>T</u> C <u>T</u>	957

LOCUS	AR178256	1493 bp	DNA	linear	PAT 20-APR-2002
DEFINITION	Sequence 38 from patent US 6319502.				
ACCESSION	AR178256				
VERSION	AR178256.1	GI:20219394			
KEYWORDS	.				
SOURCE	unknown.				
ORGANISM	unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1493)				
TITLE	Guerin-Marchand,C. and Druilhe,P.				
JOURNAL	Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes				
FEATURES	Patent: US 6319502-A 38 20-NOV-2001;				
source	Location/Qualifiers				
	1..1493				
ORIGIN	/organism="unknown"				
	/mol_type="unassigned DNA"				
Alignment Scores:					
Pred. No.:	2.1e-50	Length:	1493		
Score:	603.00	Matches:	117		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		

	US-09-837-344-20 (1-117) x AR178256 (1-1493)
QY	1 LeuGInGInGInGInGInaAGAspLeuGInGInaArgysA1aAspTrLysAsnLeuGlu 20
Db	607 TTGCAGAGCAGCAGAGAGATTATGAAcAAAGAGAGCGTGATACGAAAAAATTATGAA 666
QY	21 ArgTysLysGInuHISG1yAspL1eLeuA1gLuAspLeuTrYG1yArgLeuGlu1lePro 40
Db	667 AGAAAAAAGAAcATCGAGATATATTGACAGAGGATTATATGTGTCGTTTGAATAACCA 726
QY	41 A1a1leGluLeuProSerGluAsnGluAArgLYTyTrYLleProHISGInSerSerLeu 60
Db	727 GCTATAGAACTTCATCATCAAGAAAGAAcGTGTGATATTATATACCAcATCACTCTTTTA 786
QY	61 ProGlnAspAsnAcArgG1yAsnSerArgAspSerYsgL1eSer1le1leG1uLysThr 80
Db	787 CCTCAGAGCAACAGAGGAAATAGTAGATTCACAGAGAAATCTATATATGAAAAACA 846
QY	81 AsnArgGInSer1leThrThraAsnA1gLuG1yAArgAsp1leH1sLysG1yHISleu 100
Db	847 AATAGAGATCTTATTACAAcAAAGTTGGAAGGAGCGATATATACATTAAGGACATCTT 906
QY	101 G1uG1uLysLysAspG1ySer1leLysProGInGInLysG1uAspLysSer 117
Db	907 GAAGAAAAAGAAAGATGTTCAATATAAACCAAGACAAHAAGAAGATTAATCT 957
RESULT 9	
PRISA13	
LOCUS	1496 bp DNA linear INV 05-SEP-1994
DEFINITION	P.falciparum LSA-1 gene for liver stage antigen-1 (3').
ACCESSION	P.falciparum LSA-1
VERSION	Z30320
KEYWORDS	Z30320.1 GI:510183
SOURCE	liver stage antigen-1; LSA-1 gene.
ORGANISM	Plasmodium falciparum (malaria parasite P. falciparum)
REFERENCE	Plasmodium falciparum
AUTHORS	Eukaryota: Alveolata; Apicomplexa; Haemosporida; Plasmodium.
TITLE	1 (bases 1 to 1496)
JOURNAL	Fidock,D.A., Gras-Masse,H., Leper,J., Brahimi,K., Benmohamed,L.,
REFERENCE	Mellouk,S., Guerlin-Marchand,C., Londono,A., Raharimalala,L.,
AUTHORS	Mets,J.F., Langeley,G., Rousselhon,C., Tauxar,A., and Drulline,P.
	The Plasmodium falciparum liver stage antigen LSA-1 is well
	conserved and harbors major B- and T-cell epitopes
	J. Immunol. (1994) in press
	2 (bases 1 to 1496)
	Fidock,D.A., Gras-Masse,H., Leper,J., Brahimi,K., Benmohamed,L.,

LOCUS	DEFINITION	SEQUENCE	FROM PATENT	US 6270771.1	1482 bp	DNA	linear	PAT 17-OCT-2001
AR162974	Sequence 45	from patent	US 6270771.1					
AR162974	GI:16233440							
UNKNOWN	Unknown							
ORGANISM	Unclassified							
REFERENCE	1 (bases 1 to 1482)							
AUTHORS	Guertin-Marchand, C. and Druille, P.							
TITLE	Peptide sequences specific for the hepatic stages of <i>P. falciparum</i> bearing epitopes capable of stimulating the T lymphocytes							
JOURNAL	Patent: US 6270771-A 45 07-AUG-2001;							
FEATURES	Location/Qualifiers							
SOURCE	1..1482							
ORIGIN	1..1482							
Alignment Scores:								
Pred. No.:	4.14e-50	Length:	1482					
Score:	600.00	Matches:	116					
Percent Similarity:	100.00%	Conservative:	1					
Best Local Similarity:	99.15%	Mismatches:	0					
Query Match:	99.50%	Indels:	0					
DB:	6	Gaps:	0					
US-09-837-344-20 (1-117) x AR162974 (1-1482)								
QY	1 LeuGInGInGInGInaRgAspLeuGInGInaRgLySaLaAspThrLySAsnLeuGIn	20						
DB	610 TTGCAGAGCGACGAAAGATTGAGACAAAGAAAGCTGATACGAAAAAATTTGAA	669						
QY	21 ArgLySbLySGIuHbGlyAspIleLeuLaGlyAspLeuLyTgLyAglLeuGInIlePro	40						
DB	670 AGAAAAAGAAACATGAGATATATATACAGAGGATTTATATGCGTTTGAATACCA	729						
QY	41 AlaIleGInLeuProSerGInaAsnGInuRgLyTyrTyrIleProGInSerSerIeu	60						
DB	730 GCTATAGAACTTCATCTCAACAAATATACAGTGGATATATATACACATCAATCTTCTTA	789						
QY	61 ProGInaPAsnArgLyAsnSerArgAspSerLySgIuIleSerIleIleGInuLyThr	80						
DB	790 CCTCAGACACACAGAGGAGATAGTAGAGTTTCAAGAAATGCTATATATGAAAAACA	849						
QY	81 AsnArgGInuSerIleThrAsnValGInGlyATGATGAspIleHisLyGlyHisLeu	100						
DB	850 AATAGAGAACTCTATTACAAACAAATCTTGAGAGACAAAGGATATATCATTAAGGACATCTT	909						
QY	101 GInGInuLySbLyAspGlySerIleLySProGInuGInuLySgIuAspLySer	117						
DB	910 GAAGAAAGAAAGATGTTCAATTAACCCAGACAAAGAAAGATTAATCT	960						
RESULT 13								
LOCUS	AR178263	1482 bp	DNA	linear				
DEFINITION	Sequence 45 from patent US 6319502.							
ACCESSION	AR178263							
VERSION	AR178263.1							
KEYWORDS	GI:20219401							
SOURCE	Unknown.							
ORGANISM	Unclassified.							
REFERENCE	1 (bases 1 to 1482)							
AUTHORS	Guertin-Marchand, C. and Druille, P.							
TITLE	Peptide sequences specific for the hepatic stages of <i>P. falciparum</i>							

JOURNAL FEATURES		Location/Qualifiers	
source		1. .1482	
ORIGIN		/organism="unknown"	
		/mol_type="unassigned DNA"	
bearing epitopes capable of stimulating the T lymphocytes			
Patent: US 6319502-A 45-20-NOV-2001;			
Location/Qualifiers			
1. .1482			
/organism="unknown"			
/mol_type="unassigned DNA"			
Alignment Scores:			
Pred. No.:	4.14e-50	Length:	1482
Score:	600.00	Matches:	116
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.15%	Mismatches:	0
Query Match:	99.50%	Indels:	0
DB:	6	Gaps:	0
US-09-837-344-20 (1-117) x AR178263 (1-1482)			
QY	1	LEUGINLUGINGINAAGAPLEUGUGHARGLYALASPTHLRLYSLSAENLEUGU	20
DB	610	TTGGAAAGACGACAAAGATTTAGAAACAAGAGGCGTATACGAAAAATTTAGAA	669
QY	21	ARGYSLSYSGIWHISGLIYAAPLILEUVALGLIUAAPLEUTYRLIYARGLEUGIUIEPRO	40
DB	670	AGAAAAAGACACATGAGATATATTAGCAGAGAGATTATATGTCGTTTAGAAATACCA	729
QY	41	ALALIEGLULEUPROSERGIUAENGLIYAGLYTYTYTTLERPROHISGLINSEULEU	60
DB	730	GCTATAGAACCTTCATCAGAAAAATGAACGTCGATATATTATACCATCATCTCTTTA	789
QY	61	PROGLINAPENARAGGIYAENSERARGASPSERISGLIIESERIIEIIEGLIYLYTHR	80
DB	790	CCTCAGACCAACAGAGGAATAGTAGAGATTCCAGGAATATCTATATAGAAAAACA	849
QY	81	ASPARGLISERILETHRTTHRAENVAIGLIGLYARGARGAPLIEHISGLYHISLEU	100
DB	850	AATGAGAGATTATTACACAAAGTTGGAAGACGAAGGATATACATAAGACATCTT	909
QY	101	GIUGIUIYSLYSAAPGLYSERILEYSPROGLIUGIULYSLIUAAPLYSSEI	117
DB	910	GAAGAAAAAGAAAGATGTTCAATTAACCAAGAACAAAAAGATTAATCT	960
RESULT 14	AE014834	253001 bp	DNA linear INV 07-OCT-2002
LOCUS	Plasmodium falciparum 3D7 chromosome 10 section 6 of 7 of the		
DEFINITION	complete sequence.		
ACCESSION	AE014834 AE014185		
VERSION	AE014834.1 GI:23495173		
KEYWORDS			
ORGANISM	Plasmodium falciparum 3D7		
REFERENCE	Plasmodium falciparum 3D7		
AUTHORS	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
	1 (bases 1 to 253001)		
	Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,		
	Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,		
	James,K., Eichen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,		
	Kyes,S., Chen,M.-S., Nene,V., Shalimov,S.J., Suh,B., Peterson,J.,		
	Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D.,		
	Mather,M.W., Vaidya,A.B., Martin,D.M.A., Patilamb,A.H.,		
	Prunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,		
	Cummings,J.M., Subramanian,G.M., Mungall,C., Venter,J.C.,		
	Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and		
	Barrell,B.		
TITLE	Genome sequence of the human malaria parasite Plasmodium falciparum		
JOURNAL	Nature 419 (6906), 498-511 (2002)		
PUBMED	12368864		
REFERENCE	2 (bases 1 to 253001)		
AUTHORS	Gardner,M.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-SEP-2002) The Institute for Genomic Research, 9712		
	Medical Center Dr., Rockville, MD 20850, USA		
FEATURES	Location/Qualifiers		

source	1..253001 /organism="Plasmodium falciparum 3D7" /mol_type="genomic DNA" /isolate="3D7" /db_xref="taxon:36329" /chromosome="10" complement(79..107) /note="AT-rich"
repeat_region	/rpt_type=tandem complement(113..136) /rpt_type=tandem /rpt_unit="(a)n" 292..441
repeat_region	/rpt_type=tandem complement(292..441) /rpt_type=tandem /rpt_unit="(ta)n" 455_479
repeat_region	/rpt_type=tandem complement(455..479) /rpt_type=tandem /rpt_unit="(taa)n" 498..3257
Gene	/locus_tag="PF10_0307" 498..3257
mRNA	/locus_tag="PF10_0307" 498..3257
CDS	/locus_tag="PF10_0307" /codon_start=1 /product="hypothetical protein" /protein_id="AAJ35504.1" /db_xref="GI:23495174" /translation="MASCQKVPSSGKEKGIQLESVLENAFDELISEPKKFLHKI ITVLEBAKNEMLDIMERRRKFLPKAYENDQREDEBEYKIKNVYEKKIEGVEN VGSVENREKIMIEEKKKKKKKKSSSSNDSTGVYKSIIDENKISLKKKEKKE LKSVEENREKIMYNNKKYKNGIITHYENKEYNINLDGKYNKNEESMEPRKLYKTFE HDKNKCKDMYLYPOESHFKYSERIENLYIKSNLREKSTYNNLYDKDNENQHFK ENHTDEDTDEQNYEKRYVVKLKNVQNNHTQMDKYQINEKNNISKLKQYNNNDNN NN FYSPRSRI.FPTEPFSYASDEHNEKSDHLYMSPKDKCNINIGIILNNDKHCNNY DMSSKDKHCNNIYDMSPKDKHCNNIYIGIILNNDKHCNNIYGMSPKDKHCNNIYGI LNRKDKHCNNIYDMSPKDKHCNNIYEMSPKDKHCNNIYMSPKDKHCNNIYDMSH LSRSRANKLPSPNNIITHEKQIINENINISPOKINDLCKEKNNNECKEKKFLKSK QSPSPKSDKDDYVRGKNMVEQDDLKNMSEQRYKVEDPSYQLNRKLSNEKEIILK RNNNNL.FPFIIVIGRGRKNMGPECKDCKSPYNEIYSDDEGKCVKHPFMKNEQLKKK KININDVRSKSGSTDINSVDKNNVDRKNNANNINNADINNADINNADINNADIN NINNNNDINNADINNADINNADINNADINNADINNADINNADINNADINNADIN EKAASNEKKNGKMLNNDNINKENDEKEKEKKKKKIYTPSRHYNHNKINDSPKPK FDFPK"
repeat_region	781..821 /rpt_type=tandem 841..870
repeat_region	/rpt_type=tandem /rpt_unit="(a)n" 932..1111
repeat_region	/rpt_type=tandem /rpt_unit="(taaaa)n" 1156..1175
repeat_region	/rpt_type=tandem /rpt_unit="(taaaaa)n" complement(1369..1397) /note="AT-rich"
repeat_region	/rpt_type=tandem 1439..1601
repeat_region	/rpt_type=tandem /rpt_unit="(taa)n" 1474..1579
repeat_region	/rpt_type=tandem 1479..1588
repeat_region	/rpt_type=tandem complement(1618..1669) /note="AT-rich"
repeat_region	/rpt_type=tandem

repeat_region	1756.._2184	/rpt_type=tandem
repeat_region	1775.._1954	/rpt_type=tandem
repeat_region		/rpt_unit="(taaa)n"
repeat_region		complement(12707.._2735)
repeat_region		/note="Ar_rich"
repeat_region	2772.._2814	/rpt_type=tandem
repeat_region	2790.._2965	/rpt_type=tandem
repeat_region	2808.._2866	/rpt_type=tandem
repeat_region	3142.._3171	/rpt_type=tandem
repeat_region		/rpt_unit="(a)n"
repeat_region		complement(12774.._3347)
repeat_region		/note="Ar_rich"
repeat_region		/rpt_type=tandem
repeat_region		complement(1391.._3437)
repeat_region		/rpt_type=tandem
repeat_region		/rpt_unit="(ta)n"
repeat_region		complement(1439.._3503)
repeat_region		/rpt_type=tandem
repeat_region		/rpt_unit="(a)n"
repeat_region	3439.._3500	
repeat_region		/rpt_type=tandem
repeat_region	3440.._3500	
repeat_region		complement(13561.._3683)
repeat_region		/rpt_type=tandem
repeat_region		/rpt_unit="(ataaa)n"
repeat_region		complement(1710.._3745)
repeat_region		/note="Ar_rich"
repeat_region		/rpt_type=tandem
repeat_region		complement(13763.._3824)
repeat_region		/rpt_type=tandem
repeat_region		/rpt_unit="(ta)n"
repeat_region	3763.._3811	
repeat_region		/rpt_type=tandem
repeat_region		complement(13871.._4025)
repeat_region		/note="Ar_rich"
repeat_region		/rpt_type=tandem
repeat_region		complement(13922.._3987)
repeat_region		/rpt_type=tandem
repeat_region		/rpt_unit="(ta)n"
repeat_region	3926.._3987	
repeat_region		/rpt_type=tandem
repeat_region		complement(14037.._4082)
repeat_region		/rpt_type=tandem
repeat_region		/rpt_unit="(ataaa)n"
repeat_region	4043.._4082	
repeat_region		/rpt_type=tandem
repeat_region	4043.._4080	
repeat_region		complement(14084.._4104)
repeat_region		/rpt_type=tandem
repeat_region		/rpt_unit="(a)n"
repeat_region		complement(14135.._4181)
repeat_region		/rpt_type=tandem
repeat_region		/rpt_unit="(ta)n"
repeat_region		complement(14182.._4313)
repeat_region		/note="Ar_rich"
repeat_region		/rpt_type=tandem
repeat_region	4364.._4466	
repeat_region		/rpt_type=tandem
repeat_region	4372.._4457	
repeat_region		/rpt_type=tandem
repeat_region	4377.._4472	
repeat_region		/rpt_type=tandem
repeat_region		/rpt_unit="(ataaa)n"
repeat_region		complement(14512.._4655)

	/note="AT rich"
repeat_region	/rpe_type=tandem complement(4670, .4775)
repeat_region	/rpe_type=tandem /rpe_unit="(ca)n" 4675..4768
repeat_region	/rpe_type=tandem 4870..4912
repeat_region	/rpe_type=tandem complement(4870, .4912)
repeat_region	/rpe_type=tandem /rpe_unit="(ca)n" complement(4918, .5000)
gene	/rpe_type=tandem /rpe_unit="(taaaaa)n" <5162..>8973
mRNA	/locus_tag="PF10_0308" join(<5162, .5350, 5488, .5535, 5617, .5802, 5902, .5981, 6099, .6162, 6254, .6349, 6463, .6540, 6727, .>8973)
CDS	/locus_tag="PF10_0308" join(5162, .5350, 5488, .5535, 5617, .5802, 5902, .5981, 6099, .6162, 6254, .6349, 6463, .6540, 6727, .8973) /locus_tag="PF10_0308" /codon_start=1 /product="hypothetical protein"
Alignment Scores:	
Pred. NO.:	1,79e-47 Length: 253001
Score:	597.00 Matches: 115
Percent Similarity:	99.15% Conservative: 1
Best Local Similarity:	98.29% Mismatches: 1
Query Match:	99.00% Indels: 0
DB:	Gaps: 2 Gaps: 0
US-09-837-344-20 (1-117) x AE014834 (1-253001)	
Oy	1 LeunglNgLuglNglNaArgApLeuGlGlnAqLyAlaAsPThrLyLyAsnLeuGlu 20
Db	180455 TTGCAGAGCGACCAAGAAAGATTGTGAACAAGAAGGGCTGATTCGAAAAAATTTAGAA 180514
Oy	21 ArgLYbLYbGLNHISGLYAsPrlleuAlAGLUAsPLeuTYrgLYrGLueGluilePro 40
Db	180515 AGAAAAAGAACATGGAGATGTATTATTCAGCAGAGATTTATATGTCGTTTACAAAATACC 180574
Oy	41 AlailegIuleuProSerGIuAsngIuaRgiLYTyriLEProHIsenSerleu 60
Db	180575 GGTATAGAACTTCATCAGAAATGAACGGATTAATATATACCATCAATCTCTTAA 180634
Oy	61 ProGInaPnaPnaRgiLYaenserARgaPserLYeGluileSerlleileGluLYerThr 80
Db	180635 CCTCAGGACCAACAGAGGGAATGTAGAGATTCGAAAGAAATATCTATAATAGAAAAATACA 180694
Oy	81 AsnaRGJuseRtleThrZhenVAlGUGLYArGaRAsPrllleLYbGLYHlsleu 100
Db	180695 AATAGAGATCTATTACCACAAATGTGAAAGCGAAGGATATACATAAAGACATCTT 180754
Oy	101 GIUGLuLYbAsPGLYseRlleLYsPRoGluGlnLYbGLUAsPlySer 117
Db	180755 GAAGAAAGAAAGATCGTTCAATTAATCCAGAACAAAAGAAATATAATCT 180805
RESULT 15	
PFALSAC	843 bp DNA linear INV 19-AUG-1995
LOCUS	Plasmodium falciparum (clone KEN1-21b, KEN1-40h, KEN1-56b,
DEFINITION	PNGI-28a, PNGI-28c, BRA1-19c) liver stage-specific antigen 1
ACCESSION	(LSA-1) gene, 3' end of cds.
VERSION	L40885.1 GI:725278
KEYWORDS	exocytoxic; liver stage-specific antigen; liver stage-specific antigen 1; malaria.
SOURCE	Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE	1 (bases 1 to 843)
AUTHORS	Yang, C., Shi, Y. P., Udhayakumar, V., Alpers, M. P., Povoa, M. M., Hawley, W. A., Collins, W. E. and Lal, A. A.
TITLE	Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates
JOURNAL	Mol. Biochem. Parasitol. 71 (2), 291-294 (1995)
PUBMED	7477115
COMMENT	Original source text: Plasmodium falciparum (individual isolate Kenyan) (clone: KEN1-56b) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG1-28a) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG1-28c) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN1-21b) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN1-40h) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: BR1-19e) blood stage parasite DNA. Clone name abbreviations: KEN for Kenya; PNG for Papua New Guinea and BRA for Brazil. For each clone, the number following the KEN, PNG or BRA represents group number, the number following the '-', denotes the patient's blood sample number and the lower case letter represents recombinant clones from the corresponding isolates. Location/Qualifiers 1..843 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /isolate="Kenyan" /db_xref="taxon:5833" /clone="KEN1-56b" /clone="PNG1-28a" /clone="PNG1-28c" /clone="KEN1-21b" /clone="KEN1-40h" /clone="BR1-19e" /dev_stage="blood stage parasite" 1..843 /gene="LSA-1" <1..843 /gene="LSA-1" /codon_start=1 /product="liver stage-specific antigen 1" /protein_id="AA859231.1" /db_xref="GI:950664" /translation="RKADITKKNLEKKKHHGDDILAEDYGRLEPALELPBENEGYIT /PHOSLPQDNKGRSRSKSIIEKTRRESITTNVEGRRIHKGHLEKKDGSIKPEQ KYEDSADIQNHLETETVNI SPVDFQISKEYDELSAEDYDGLIDEBDDDLDEKPIV QYDFDQBNENIGYKELEDLIEKKENMLDDDEGIEKSEFELSEKIKKGKYEYTKDN NPKNDKSLYDENHKYKYKNDKQYNNKEKERIKSLFHPFDQDNEIQVDELSIDITKY PMKL"
ORIGIN	228 /gene="LSA-1" /note="clone PNG1-28a" /replace="t"
variation	
Alignment Scores:	
Pred. No.:	8,45e-46 Length: 843
Score:	554.00 Matches: 107
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	91.87% Gaps: 0
DB:	2
US-09-837-344-20 (1-117) x PFALSAC (1-843)	
QY	11 ArgLySaLaApPTrLrLyLySaLnLeuGluArgLyLyGluLHtGlyAaPllLeuLa 30
Db	1 AGGAAGCGTGATACGAAAAAAATTGTGAAAGAAAAAGAACATGAGATATATTGCA 60
QY	31 GLuApLeuTyrGlyVtGLeuGluLeProAlaIleGluLeuProSerGluAsnGluArg 50
Db	61 GAGGATTATATAGGCGTTTGAATAATCAACACTTATGAACTTCATCAGAAAAAGAACT 120

[illegible]

Search completed: December 7, 2005, 10:31:35
Job time : 2652.64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:22:20 ; Search time 306.146 Seconds
(without alignments)
2547.054 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LQEQKRDLEQRKADTKKLE.....GHLEKKDSIKPEKEDKS 117

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1h
-Q/cgnc2 1/USPTO.spool.h/US09837344/runat.05122005.133531.726/app.query.fasta_1.718
-DB=an.GeneSeq -OFFT=fatgap -SUFFIX=ring -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCAL=GN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09837344 @CGC 1.1 900 @runat.05122005.133531.726 -NCPU=6 -ICPU=3
-NO MMAP -LARGEJITTER -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N_GeneSeq_21:*
1: geneSeqn1980s:*
2: geneSeqn1990s:*
3: geneSeqn2000s:*
4: geneSeqn2001a:*
5: geneSeqn2001b:*
6: geneSeqn2002a:*
7: geneSeqn2002b:*
8: geneSeqn2003a:*
9: geneSeqn2003b:*
10: geneSeqn2003c:*
11: geneSeqn2003d:*
12: geneSeqn2004a:*
13: geneSeqn2004b:*
14: geneSeqn2005a:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	603	100.0	950	2	AAQ28115
2	603	100.0	1496	2	AAQ28119
3	602	99.8	1371	12	ADO21941
4	602	99.8	1374	12	ADO21919

5	564	93.5	1320	2	AAQ80916	AAQ80916 Plasmodium
6	94	15.6	1384	14	ADZ40618	AdZ40618 PFCITL/HTL
7	94	15.6	1385	9	ADA49438	Ada49438 Muli-1-epi
8	94	15.6	1385	12	ADO24116	ADO24116 Epigene C
9	90.5	15.0	969	5	AA575358	AA575358 DNA encod
10	90.5	15.0	969	5	AA571680	AA571680 DNA encod
11	90.5	15.0	969	5	AA592259	AA592259 DNA encod
12	90.5	15.0	4599	2	AAT72052	AAT72052 RRP3 telo
13	90	14.9	4080	13	ADR25659	ADR25659 Breast ca
14	89.5	14.8	696	4	AAH84231	AAH84231 Human cel
15	89.5	14.8	699	4	AAH84230	AAH84230 Human cel
16	89.5	14.8	717	4	AAH84229	AAH84229 Human cel
17	89.5	14.8	774	4	AAH84228	AAH84228 Human cel
18	89.5	14.8	819	4	AAH84227	AAH84227 Human cel
19	89.5	14.8	1669	4	AAH84226	AAH84226 Human cel
20	87.5	14.5	700	2	AAV88907	AAV88907 EST clone
21	87.5	14.5	1391	3	AAZ57866	AAZ57866 Protein T
22	87.5	14.5	2554	6	ABA92314	ABA92314 Human cyl
23	87	14.4	903	13	ADS51252	ADS51252 Bacterial
24	87	14.4	2175	12	ADN04691	ADN04691 Antipsoxi
25	86.5	14.3	601	13	ACN45782	ACN45782 Cotton pr
26	86.5	14.3	3684	4	ABL23583	ABL23583 Drosophi
27	86	14.3	1344	4	AAH53753	AAH53753 S. epider
28	86	14.3	1353	4	AAH53916	AAH53916 S. epider
29	86	14.3	3001	4	AAH54449	AAH54449 S. epider
30	86	14.3	3974	9	AAH54481	AAH54481 S. epider
31	85.5	14.2	2298	9	ADA02680	ADA02680 Human TOP
32	85.5	14.2	2298	10	ADB72418	ADB72418 Human TIO
33	85.5	14.2	2298	10	ADE95928	ADE95928 Human TOP
34	85.5	14.2	2469	6	ABR84206	ABR84206 Human CDN
35	85.5	14.2	3440	12	AD116295	AD116295 Human nuc
36	85.5	14.2	3477	6	ABL70013	ABL70013 Pancreas
37	85.5	14.2	3477	6	ABK84207	ABK84207 Human CDN
38	85.5	14.2	3734	12	ADL82892	ADL82892 Human PRO
39	85.5	14.2	3734	13	ADL82892	ADL82892 Human PRO
40	85.5	14.2	3729	8	ABZ23989	ABZ23989 Human top
41	85.5	14.2	3740	8	ACF62748	ACF62748 Cancer ba
42	85.5	14.2	3740	8	ADB20867	ADB20867 MRP1 base
43	85.5	14.2	3740	9	ADA02679	ADA02679 Human TOP
44	85.5	14.2	3740	10	ADB87956	ADB87956 Human UGT
45	85.5	14.2	3740	10	ADB72417	ADB72417 Human TIO

ALIGNMENTS

RESULT 1
AAQ28115
ID AAQ28115 standard; DNA; 950 BP.
XX
AC AAQ28115;
XX
DT 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX
DE P. falciparum LSA-R-NR coding sequence.
XX
KW Malaria; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;
KW paludism; liver stage-specific antigen; ss.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FH repeat_region 3..629
FT /*tag= a
FT /rpt type= TANDEM
FT repeat_unit 3..53
FT /*tag= b
XX
XX W09213884-A1.
XX
XX 20-AUG-1992.
XX
XX 05-FEB-1992; 92MO-FR000104.

```
XX 05-FEB-1991; 91FR-00001286.
XX (INSP ) INST PASTEUR.
XX Guerimarchand C, Druilhe P;
XX WPI, 1992-299985/36.
XX P-PSDB; AAR26941.
XX
XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX vaccination against, treatment of and diagnosis of malaria.
XX
XX Claim 17; Fig 2; 81pp; French.
XX
XX A genomic DNA bank of P. falciparum EcoRI fragments prepared in lambda
XX gII1 was used to transform E.coli. The expression library was screened
XX with human antisera against antigens of all stages of P. falciparum. The
XX library was rescreened with antibodies affinity- purified on a clone
XX which was able to recognise antibodies specific to the hepatic phase.
XX About 40 clones were detected which produced a characteristic USA
XX epitope. The clone with the largest insert (950 bases) encoded USA-R-NR
XX containing a 12-repeat region followed by a non-repeat region. Preferred
XX antigenic polypeptides of the invention are derived from the amino acid
XX sequence of USA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 950 BP; 456 A; 127 C; 206 G; 161 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 2,996-63 Length: 950
XX Score: 603.00 Matches: 117
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-09-837-344-20 (1-117) x AAQ28115 (1-950)
XX
XX 1 LeuGInGInGInGInAArgAspLeuGInAArgLySA1aAspThrLySAAsnLeuGlu 20
XX 600 TTGCAAGAGCAGCAAGAGATTGTAACCAAGAAAGCGTGATACGAAAAAATTTAGAA 659
XX
XX 21 ArgLySLySGuLH1eG1yAspL1eLeuA1aGluAspLeuTyG1yArgLeuGlu1lePro 40
XX 660 AGAAAAAGAAACAATGAGATATATTAGCAGAGATTTATATGCTCTTGAATATACCA 719
XX
XX 41 Ala1leGluLeuProSerG1uAsnG1uAArgG1yTyTyT1leProH1eGInSerSerleu 60
XX 720 GCTATAGAACTTCATCAGAAAAATGACGTGATATTATATACCATCAATCTCTTTA 779
XX
XX 61 ProG1naAspAenAArgG1yAenSerAArgAspSerLySGu1leSer1le1leG1uLyThr 80
XX 780 CCTCGAGCAACAGAGGAAATAGTAGAGATTCCAGAAATATCTATATAGAAAAACA 839
XX
XX 81 AenAArgG1uSer1leThrThraenValG1uG1yAArgAspL1eH1eLySG1yH1eLeu 100
XX 840 AATAGAGAACTATATTAACAACAAATGTTGAGAGCAAGAGATATCATTAAGGACATCTT 899
XX
XX 101 G1uG1uLySLySAAspG1ySer1leLySPProG1uG1uLySG1uAspLySer 117
XX 900 GAAAGAAAGAAAGATGTTCAATTAACACAGAAACAAAAAGAGATTAATCT 950
XX
XX RESULT 2
XX AAQ28119
XX ID AAQ28119 standard; DNA; 1496 BP.
XX
XX AAQ28119;
XX AC
XX 25-MAR-2003 (revised)
XX DT 08-FEB-1993 (first entry)
XX
XX P.falciparum LSA gene 3' region.
XX
```

```
KW Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
KW paludism; liver stage-specific antigen; ss.
XX Plasmodium falciparum.
XX
XX Key Location/Qualifiers
XX CDS 1..1482
XX FT /*cag= c
XX FT /product= "LSA_C-terminal_region"
XX FT repeat_region 37..639
XX FT /*cag= a
XX FT repeat_unit 37..87
XX FT /*cag= b
XX FT /rpt_type= TANDEM
XX
XX W09213884-A1.
XX
XX 20-AUG-1992.
XX
XX 05-FEB-1992; 92WO-FR000104.
XX
XX 05-FEB-1991; 91FR-00001286.
XX (INSP ) INST PASTEUR.
XX
XX Guerimarchand C, Druilhe P;
XX
XX WPI, 1992-299985/36.
XX P-PSDB; AAR26944.
XX
XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX vaccination against, treatment of and diagnosis of malaria.
XX
XX Claim 2; Fig 8-10; 81pp; French.
XX
XX This nucleotide sequence is the 3' part of the P. falciparum liver-stage
XX specific antigen (USA) gene. It codes for a polypeptide sequence which
XX carries a T cell epitope characteristic of a protein produced in
XX hepatocytes infected with P. falciparum. The polypeptide can be used in
XX the preparation of vaccines against malaria. (Updated on 25-MAR-2003 to
XX correct PN field.)
XX
XX Sequence 1496 BP; 716 A; 169 C; 300 G; 311 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 5,46-63 Length: 1496
XX Score: 603.00 Matches: 117
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-09-837-344-20 (1-117) x AAQ28119 (1-1496)
XX
XX 1 LeuGInGInGInGInAArgAspLeuGInAArgLySA1aAspThrLySAAsnLeuGlu 20
XX 610 TTGCAAGAGCAGCAAGAGATTGTAACCAAGAAAGCGTGATACGAAAAAATTTAGAA 669
XX
XX 21 ArgLySLySGuLH1eG1yAspL1eLeuA1aGluAspLeuTyG1yArgLeuGlu1lePro 40
XX 670 AGAAAAAGAAACAATGAGATATATTAGCAGAGATTTATATGCTCTTGAATATACCA 729
XX
XX 41 Ala1leGluLeuProSerG1uAsnG1uAArgG1yTyTyT1leProH1eGInSerSerleu 60
XX 730 GCTATAGAACTTCATCAGAAAAATGAAAGTGATATTATATACCATCAATCTCTTTA 789
XX
XX 61 ProG1naAspAenAArgG1yAenSerAArgAspSerLySGu1leSer1le1leG1uLyThr 80
XX 790 CCTCGAGCAACAGAGGAAATAGTAGAGATTCCAGAAATATCTATATAGAAAAACA 849
XX
XX 81 AenAArgG1uSer1leThrThraenValG1uG1yAArgAspL1eH1eLySG1yH1eLeu 100
XX 850 AATAGAGAACTATTAACAACAAATGTTGAGAGCAAGAGGATATATACATTAAGGACATCTT 909
XX
```


QY 101 GUGLULYLYEAAPGlySerIleYsPProGluGlnlyGluAspIysSer 117
 DB 910 GAAGAAAAGAAAGATGGTCAATPAACCAAGAACAAAGATTAATCT 960

RESULT 3

ADO21941
 ID ADO21941 standard; DNA; 1371 BP.

AC ADO21941;

DT 12-AUG-2004 (first entry)

DE USA-NRC(H) construct DNA derived from Malaria parasite USA-1.

XX USA-NRC; liver stage antigen; USA-1; T-cell; B-cell epitope;
 KM parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
 XX malaria parasite P. falciparum; USA-NRC(H) construct; db; gene.

OS Plasmodium falciparum.
 XX Synthetic.

XX Key Location/Qualifiers
 FH 1. .1371
 FT /*tag= a
 FT /product= "USA-NRC(H) construct protein derived from
 Malaria parasite USA-1"

PN MO2004044167-A2.

PD 27-MAY-2004.

PF 12-NOV-2003; 2003MO-US036011.

PR 12-NOV-2002; 2002US-0425719P.

PA (REED-) REED ARMY INST RES WALTER.

PI Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
 PI Barbosa A;

XX WPI; 2004-420309/39.

DR P-PSDB; ADO21942.

XX Recombinant USA-NRC polypeptide for use as a vaccine and a diagnostic
 PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
 PT (USA-1) epitope.

PS Claim 16; SEQ ID NO 25; 90pp; English.

XX The invention relates to a novel recombinant USA-NRC polypeptide
 CC comprising liver stage antigen (USA-1) T-cell and B-cell epitopes. USA-1
 CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
 CC falciparum. The polypeptide of the invention demonstrates antimalarial
 CC activities and may be useful as a vaccine during immunotherapy and as a
 CC diagnostic agent, preferably for diagnosing malaria. The current sequence
 CC is that of the USA-NRC(H)Mut construct DNA of the invention. The
 CC construct is harmonized for expression in Escherichia coli and comprises
 CC the harmonized N-terminal, C-terminal and 2 tandem repeats of the Malaria
 CC parasite P. falciparum USA-1 in addition to a C-terminal His6 tag.

XX Sequence 1371 BP; 539 A; 358 C; 283 G; 191 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.37e-63 Length: 1371
 Score: 602.00 Matches: 116
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.15% Mismatches: 0
 Query Match: 99.83% Indels: 0
 DB: 12 Gaps: 0

US-09-837-344-20 (1-117) x ADO21941 (1-1371)

QY 1 LeuGInGluGInAArgAAspleuGInArgLySAlaAspThrLySAsnLeuGIn 20
 DB 463 CTCGAGAGCAACAGCGGCACTCGAACAGCCCAAGCTGACACGAAAAAACTGGAA 522
 QY 21 ArgLySLeuGInHiSgIyAspIleLeuAlaGluAspleuTyrGlyArgLeuGInIlePro 40
 DB 523 CGCAAAAAGAAACACGCGCACTTCCTGCTGAGGACCTGTAGCGCGCTGGAAAACTCCA 582
 QY 41 AlaIleGluLeuProSerGlyAsnGluArgGlyTyrTyrIleProHisGlnSerSerLeu 60
 DB 583 GCTATCAACTCCCATCCCAAAACGAACCGGCTACTACATCCACACAGACGACCTG 642
 QY 61 ProGlnAspAsnArgGlyAsnSerArgAspSerLySgIuIleSerIleIleGlyThr 80
 DB 643 CCACAAAGATTAATCGCGGAACTCCCGCAAGTAAAGAAATCAGATCATCGAAAAAAC 702
 QY 81 AsnArgGlySerIleThrThrAsnValGluGlyArgArgAspIleHisLySgIyHisLeu 100
 DB 703 AACCGGAAAGCATTCACCAACGTCGAAGGCCCGCGACATCCCAAAAGGCCACCTC 762

RESULT 4

ADO21919
 ID ADO21919 standard; DNA; 1374 BP.

AC ADO21919;

DT 12-AUG-2004 (first entry)

DE USA-NRC(H)Mut construct DNA derived from Malaria parasite USA-1.

XX USA-NRC; liver stage antigen; USA-1; T-cell; B-cell epitope;
 KM parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
 KM malaria parasite P. falciparum; USA-NRC(H)Mut construct; db; gene;
 XX mutant.

XX Plasmodium falciparum.

OS Synthetic.

XX Key Location/Qualifiers
 FH 1. .1374
 FT /*tag= a
 FT /product= "USA-NRC(H)Mut construct protein derived from
 Malaria parasite USA-1"

PN MO2004044167-A2.

PD 27-MAY-2004.

PF 12-NOV-2003; 2003MO-US036011.

PR 12-NOV-2002; 2002US-0425719P.

PA (REED-) REED ARMY INST RES WALTER.

PI Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
 PI Barbosa A;

XX WPI; 2004-420309/39.

DR P-PSDB; ADO21920.

XX Recombinant USA-NRC polypeptide for use as a vaccine and a diagnostic
 PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
 PT (USA-1) epitope.

PS Claim 17; SEQ ID NO 3; 90pp; English.

XX The invention relates to a novel recombinant USA-NRC polypeptide
 CC comprising liver stage antigen (USA-1) T-cell and B-cell epitopes. USA-1
 CC protein is found within the parasitophorous vacuole (PV) of Plasmodium

CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct DNA of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to an amino acid insertion
CC within the T5 P. falciparum epitope and a C-terminal His6 tag.

XX Sequence 1374 BP; 539 A; 358 C; 285 G; 192 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6,39e-63	Length:	1374
Score:	602.00	Matches:	116
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.15%	Mismatches:	0
Query Match:	99.83%	Indels:	0
DB:	12	Gaps:	0

US-09-837-344-20 (1-117) x ADO21919 (1-1374)

```
QY      1  LeuGInGInGInGInAArgAspLeuGInAArgLysAlaAspThrLysLeuAsnLeuGlu 20
        |||||
DB      463 CTCCAGAGCAACACGCGACCTTGAGAACGCGCAAGCTTGACACGAAAAAAACCTGGAA 522

QY      21  ArgLysLysGluHisGlyAspIleLeuAlaGluAspLeuTyrgLysArgLeuGluIlePro 40
        |||||
DB      523 CCGCAAAAAGAAACACGCGACCTTCTGCTGAGACCTGTACGGCCGCTGAAATCCCA 582

QY      41  AlaIleGluLeuProSerGluAsnGluArgGlyTyrgLysIleProHisGlnSerSerLeu 60
        |||||
DB      583 GGTATCGAAGCTCCATCCGAAACGAAACGCGGCTACTACATCCACACGACGACGCTG 642

QY      61  ProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThr 80
        |||||
DB      643 CCACAAGATTAATCGCGGAACTCCGCGACAGTAAGAAATCAGCATCATCGAAAAAAC 702

QY      81  AsnArgLysSerIleThrThrAsnValGluGlyArgArgAspIleHisIleGlyHisLeu 100
        |||||
DB      703 AACCCGAAAGCATTAACCAACCAAGCTGGAAGCCCGCCGCAATCCACAAAGCCACCTC 762

QY      101  GluGluLysLysAspGlySerIleLysProGluGlnLysGluAspLysSer 117
        |||||
DB      763 GAAGAAAGAAAGACGGCTCCATCAACAGAAACGAAAGAAAGCAAAAGC 813

RESULT 5
AA080916
ID      AA080916 standard; cDNA; 1320 BP.
XX
XX
XX      AA080916;
AC
XX      16-OCT-2003 (revised)
DT      25-MAR-2003 (revised)
DT      29-AUG-1995 (first entry)
XX
XX      Plasmodium falciparum LSA-1 repeatless gene cDNA.
DE
XX      Plasmodium falciparum LSA-1 repeatless gene; recombinant poxvirus;
KW      multicomponent multistage malarial vaccines; immunogens;
KW      malaria diagnosis; ss.
XX
XX      Plasmodium falciparum; (pLSARPLS.14L.1).
OS
XX      WO9428930-A1.
PN
XX      22-DEC-1994.
XX
XX      10-JUN-1994; 94WO-US006652.
PF
XX      11-JUN-1993; 93US-00075783.
PR      09-JUN-1994; 94US-00257073.
XX
XX      (VIRO-) VIROGENETICS CORP.
```

XX
XX Paoletti E, De Taisne C, Tine JA;
XX WPI; 1995-036113/05.
DR
XX
XX Recombinant poxvirus contg. Plasmodium DNA in non-essential region -
PT useful in vaccines against malaria and for prodn. of Plasmodium
XX immunogens.

XX Claim 3; Fig 11; 183pp; English.

XX AA080916 is the P. falciparum LSA-1 repeatless gene cDNA sequence. New
CC recombinant poxviruses containing either the SERA, ABRA, Plnsp10, AMA-1,
CC Pf25, Pf26, CSP, Pf28P2, LSA-1, LSA-1 repeatless, MSA-1, MSA-1 (N-
CC terminal p83 or C-terminal gp42) genes, or a combination of these in non-
CC essential regions of their genomes are claimed. These poxviruses (pref.
CC with a virulence reducing genomic deletion or disruption) can be used as
CC vaccines against malaria and for the prodn. of Plasmodium immunogens.
CC These viruses provide multicomponent, multistage vaccines due to their
CC expression of sporozite, liver stage, blood stage and sexual stage
CC proteins. (Updated on 25-MAR-2003 to correct PN field.) (updated on 16-
CC OCT-2003 to standardise OS field)

XX Sequence 1320 BP; 619 A; 128 C; 227 G; 346 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,45e-58	Length:	1320
Score:	564.00	Matches:	110
Percent Similarity:	91.94%	Conservative:	4
Best Local Similarity:	88.71%	Mismatches:	2
Query Match:	93.53%	Indels:	8
DB:	2	Gaps:	1

US-09-837-344-20 (1-117) x AA080916 (1-1320)

```
QY      2  GInGInGInGInAArgAspLeuGIn-----ArgLysAla 13
        |||
DB      427 GAAACAGACAAAGAGATCTTGAAAGAAAAAGCGCCGATCTAAAGAAACGAGAGAGCT 486

QY      14  AspThrLysLysAsnLeuGluArgLysLysGluHisGlyAspIleLeuAlaGluAspLeu 33
        |||||
DB      487 GATCGAAAAAAATTTAGAAAGAAAGAAAGAAACATGAGATGATTTAGCAGAGATTTA 546

QY      34  TyrGlyArgLeuGluIleProAlaIleGluLeuProSerGluAsnGluArgGlyTyrgLys 53
        |||||
DB      547 TATGTCGTTTAGAAATTAACGCTATAGAACTTCATCAGAAATGAACGCGATATTAT 606

QY      54  IleProHisGlnSerSerLeuProGlnAspAsnArgGlyAsnSerArgAspSerLysGlu 73
        |||||
DB      607 ATACCAATCAATCTTCTTTACTCAGACAAACAGAGGAATAGTAGATTCCAAAGAA 666

QY      74  IleSerIleIleGluLysThrAsnArgGluSerIleThrThrAsnValGluGlyArgArg 93
        |||||
DB      667 ATATCTATATATGAAAAAACAATAGAAATCTATTCAACAAATGTTGAAGACGAAAG 726

QY      94  AspIleHisLysGlyHisLeuGluGluLysLysAspGlySerIleLysProGluGlnLys 113
        |||||
DB      727 GATATACCTTAAGACATCTTGAAAGAAAGAAAGATGTTCAATTAACCAAGAAACAAAA 786

QY      114  GluAspLysSer 117
        |||||
DB      787 GAAGATAAATCT 798

RESULT 6
ADZ40618
ID      ADZ40618 standard; DNA; 1384 BP.
XX
XX      ADZ40618;
AC
XX      30-JUN-2005 (first entry)
DT
XX      PfcTL/HTL/N multi-epitope construct encoding DNA.
DE
XX
```

KW multi-epitope; immune stimulation; HIV infection; anti-HIV;
KW immunostimulant; vaccine; immunogenicity; gene; ds.
XX Synthetic.
XX OS
PN WO2005033265-A2.
XX
PD 14-APR-2005.
XX
PD 26-APR-2004; 2004WO-US012732.
XX
PF 25-APR-2003; 2003US-0465229P.
XX
PR (EPIM-) EPIMUNE INC.
XX
PA Sette A, Chesnut RW, Newman MJ, Livingston BD;
XX
PI WPI; 2005-285416/29.
XX
DR P-PSDB; AD240617.
XX
DR New multi-epitope polynucleotide, useful for designing vaccines that
XX provide increased immunogenicity, and for inducing immune responses
PT against HIV in an individual.
XX
PS Disclosure; Fig 181; 261pp; English.
XX
XX The invention relates to a novel multi-epitope polynucleotide. The
CC invention further comprises: a composition comprising the polynucleotide;
CC a cell comprising the polynucleotide; a method for inducing an immune
CC response against human immunodeficiency virus (HIV) in an individual;
CC comprising administering the polynucleotide, composition or cell to the
CC individual; and a method for making the polynucleotide, composition or
CC cell. The multi-epitope polynucleotide has anti-HIV and immunostimulant
CC activities. The polynucleotide is useful for designing vaccines that
CC provide increased immunogenicity, and for inducing an immune response
CC against HIV. This polynucleotide sequence represents an PfCTL/HTL/N multi-
CC -epitope construct encoding DNA of the invention.
XX
SQ Sequence 1384 BP; 351 A; 347 C; 361 G; 325 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0.239 Length: 1384
Score: 94.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 15.59% Indels: 0
DB: Gaps: 0
US-09-837-344-20 (1-117) x AD240618 (1-1384)
QY 40 ProhlaarlegluuPProSerGluAenGluArgGlyTyrTyrIleProHISGInSerSer 59
DB 360 CCTGGACCAAGCTGCGCTCCGAAACGAAAGAGATCTACTACCTCACCAGAGCAGC 419
QY 60 leu 60
DB 420 CTC 422
RESULT 7
ADA49438
ID ADA49438 standard; DNA; 1385 BP.
XX
AC ADA49438;
XX
DT 20-NOV-2003 (first entry)
XX
DE Multi-epitope construct #1 DNA.
XX
XX ds; gene; multi-epitope; immunogenic; epitope;
KW major histocompatibility complex; MHC class I; MHC class II;
XX functional epitope.
OS Synthetic.

OS Plasmodium falciparum.
XX
XX US2002119127-A1.
XX
XX 29-AUG-2002.
XX
XX 27-JUN-2001; 2001US-00894018.
XX
XX 28-DEC-1999; 99US-0173390P.
XX
XX 28-DEC-2000; 2000WO-US035568.
XX
XX 16-APR-2001; 2001US-0284221P.
XX
XX (SETT/) SETTE A.
XX
XX (CHES/) CHESNUT R.
XX
XX (LIV/) LIVINGSTON B D.
XX
XX (BAKE/) BAKER D M.
XX
XX (NEWM/) NEWMAN M J.
XX
XX (BROW/) BROWN D H.
XX
PI Sette A, Chesnut R, Livingston BD, Baker DM, Newman MJ, Brown DH;
XX
XX WPI; 2003-615704/58.
XX
DR P-PSDB; ADA49437.
XX
XX Designing multi-epitope construct having major histocompatibility complex
PT class I and II epitope nucleic acids, by selecting mixture of amino acid
XX insertions at junctions of construct to minimize junctional epitopes.
XX
XX Disclosure; Fig 181; 78pp; English.
XX
XX The invention relates to a method of designing multi-epitope constructs
CC comprising major histocompatibility complex (MHC) class I and II (CTL)
CC epitope nucleic acids (CEN), involves sorting CEN, introducing flanking
CC amino acid residue selected from specified amino acid residues given in
CC specification at C+1 position of CEN, introducing amino acid spacer
CC residues between two CEN, and selecting the constructs having less
CC junctional epitopes. The method is useful for designing a multi-epitope
CC construct having multiple epitope nucleic acid. The method avoids or
CC minimises the occurrence of junctional epitopes and maximises the
CC immunogenicity and/or antigenicity of multi-epitope vaccines. The present
CC sequence represents DNA encoding a multi-epitope construct.
XX
SQ Sequence 1385 BP; 351 A; 347 C; 361 G; 326 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0.239 Length: 1385
Score: 94.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 15.59% Indels: 0
DB: Gaps: 0
US-09-837-344-20 (1-117) x ADA49438 (1-1385)
QY 40 ProhlaarlegluuPProSerGluAenGluArgGlyTyrTyrIleProHISGInSerSer 59
DB 361 CCTGGACCAAGCTGCGCTCCGAAACGAAAGAGATCTACTACCTCACCAGAGCAGC 420
QY 60 leu 60
DB 421 CTC 423
RESULT 8
ADO24116
ID ADO24116 standard; DNA; 1385 BP.
XX
AC ADO24116;
XX
XX 01-JUL-2004 (first entry)
XX
XX Epigene construct PfCTL/HTL(N) DNA.
XX
XX ds; gene; hepatotropic; virucide; antiinflammatory; anti-HIV; cytostatic;

KW epitope; cytotoxic T lymphocyte; CTL; helper T lymphocyte; HTL;
KW immune response; hepatitis B virus; HBV; HIV; cancer; lymphoma.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT exon 1..1377
FT /*tag= a
XX
PN WO2004031210-A2.
XX
PD 15-APR-2004.
XX
PF 03-OCT-2003; 2003WO-US031303.
XX
PR 03-OCT-2002; 2002US-0415463P.
PR 22-OCT-2002; 2002US-0419773P.
XX
PA (EPRM-) EPIMUNE INC.
PA (GENM-) GENENCOR INT INC.
PI Sette A, Chenut R, Newman MJ, Livingston BD, Babe LM, Chen Y;
PI Deyoung LM, Huang MTF, Power SD;
XX
DR WPI; 2004-330143/30.
DR P-PSDB; ADO24115.
XX
PT New multi-epitope polynucleotides encoding cytotoxic T lymphocyte and/or
PT helper T lymphocyte epitopes, useful for inducing or stimulating an
PT antiviral or anticancer immune response, especially against hepatitis B
PT virus.
XX
PS Disclosure; SEQ ID NO 266; 401bp; English.
XX
CC The invention relates to a multi-epitope polynucleotide (I) encoding
CC cytotoxic T lymphocyte (CTL) and/or helper T lymphocyte (HTL) epitopes.
CC The composition and methods are useful for inducing or stimulating an
CC immune response against hepatitis B virus (HBV), other viruses (e.g.
CC HIV), or against cancers such as prostate cancer or lymphoma. This
CC sequence corresponds to an example of an epigene construct used in the
CC invention.
XX
SQ Sequence 1385 BP; 351 A; 347 C; 361 G; 326 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0.239 Length: 1385
Score: 94.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 15.59% Indels: 0
DB: Gaps: 0
XX
US-09-837-344-20 (1-117) x ADO24116 (1-1385)
QY 40 ProAlaIleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHisGlnSerSer 59
DB 361 CCTGACCAAGCTCTGCTCCGAAACGAAAGAGATTAATCACTCAACCAAGACAC 420
QY 60 Leu 60
DB 421 CTC 423
XX
RESULT 9
ID AAS75358/c
XX AAS75358 standard; cDNA; 969 BP.
XX
AC AAS75358;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #11162.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG11171.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 11162; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 969 BP; 175 A; 188 C; 231 G; 375 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0.399 Length: 969
Score: 90.50 Matches: 30
Percent Similarity: 47.66% Conservative: 21
Best Local Similarity: 28.04% Mismatches: 35
Query Match: 15.01% Indels: 21
DB: Gaps: 6
XX
US-09-837-344-20 (1-117) x AAS75358 (1-969)
QY 1 LeuGngInGngInGlnArgAspLeuGluGlnArgTysAlaAspThrLysAsnLeuGlu 20
DB 600 CTCACGACAGCTCCACACGACCTGACGTGAGGCTCCGACTGTTAGAAGAAAACTAAC 541
QY 21 ArgLysIlyGluHisGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGluIlePro 40
DB 540 AAACGAAAGACATCCACAG-----CAAAACCC 511
QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrTyr-----IleProHisGln 57
DB 510 ATCTGTACGTCACTATCAAAAGACCAAGATATATTCAGAGAGAACTTCCCAACT 451

```
QY 58 SerSerLeuProGln---ApsAenArgGlyAenSerArgApsSerLySGluIleSerIle 76
DB 450 AGCAAGACAGCGCCAGATTTCAGAAATACACAGAAATCCCAAGAAATA---CTC 394
QY 77 lIeGluLyThrAsnArgIuSerIleThrThAsnValGluLyArgArgApsIleHis 96
DB 393 CTCGAGAAAGAGCAAC-----TCCAGACACATTAATTGTCAG-----ATTCA 352
QY 97 LySGlyHISLeuGluGluLyS 103
DB 351 CAAGTTGAATGAAGAAAA 331

RESULT 10
AAS71680/c
ID AAS71680 standard; cDNA; 969 BP.
XX
XX AAS71680;
AC
XX 13-FEB-2002 (first entry)
DT
XX DNA encoding novel human diagnostic protein #7484.
DE
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX MO200175067-A2.
XX
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001MO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR P-PSDB; ABG07493.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 7484; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridization probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX patent sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pat_sequences
XX
XX Sequence 969 BP; 175 A; 188 C; 231 G; 375 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 0.399
Score: 90.50
Percent Similarity: 47.66%
Best Local Similarity: 28.04%
Query Match: 15.01%
DB: 5 Gaps: 6

US-09-837-344-20 (1-117) x AAS71680 (1-969)
QY 1 LeuGluGluGlnGlnArgApsLeuGluGlnArgLyAlaAspThrLySAsnLeuGlu 20
DB 600 CTCGACGAAGCTCCACACAGACCTCGACGCTGACGCTTGAAGAAACTAAC 541
QY 21 ArgLyLySGluHISGlyAspIleLeuAlaGluApsLeuTrGlyArgLeuGluIlePro 40
DB 540 AAACAGAAAAGACATCCACAG-----CAAAACCC 511
QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrTyr-----IleProHisGln 57
DB 510 ATCTGACGTCACATCATCAAGACCAAGAGATTAATTCAGAGAACTTCCCAACT 451
QY 58 SerSerLeuProGln---ApsAenArgGlyAenSerArgApsSerLySGluIleSerIle 76
DB 450 AGCAAGACAGCGCCAGATTTCAGAAATACACAGAAATCCCAAGAAATA---CTC 394
QY 77 lIeGluLyThrAsnArgIuSerIleThrThAsnValGluLyArgArgApsIleHis 96
DB 393 CTCGAGAAAGAGCAAC-----TCCAGACACATTAATTGTCAG-----ATTCA 352
QY 97 LySGlyHISLeuGluGluLyS 103
DB 351 CAAGTTGAATGAAGAAAA 331

RESULT 11
AAS92259
ID AAS92259 standard; cDNA; 969 BP.
XX
XX AAS92259;
AC
XX 13-FEB-2002 (first entry)
DT
XX DNA encoding novel human diagnostic protein #28063.
DE
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX MO200175067-A2.
XX
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001MO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR P-PSDB; ABG28072.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 28063; 103bp; English.
XX
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:30:56 ; Search time 92.9515 Seconds

(without alignments)
2237.454 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LBBQRDLBQRKADTKKLE.....GHLEKKDSIKPEKKEDKS 117

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Egapop 10.0 , Egapext 0.5	
Egapop 6.0 , Egapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q/cgcn2_1/USPTO.spool_h/US09837344/runat_05122005_133533_763/app_query.fasta_1.718
-Database=Patentc NA -QFMT=faaetap -SUFFIX=rni -MINMATCH=0.1 -LOOFCF=0
-LOOPEXT=0 -INITs=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=epct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=apto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09837344 @CGN 1.1 237 @runat_05122005_133533_763 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patentc NA.*
1: /cgcn2_6/ptodata/1/ina/1.COMB.seq.*
2: /cgcn2_6/ptodata/1/ina/5.COMB.seq.*
3: /cgcn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgcn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgcn2_6/ptodata/1/ina/H.COMB.seq.*
6: /cgcn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
7: /cgcn2_6/ptodata/1/ina/PP.COMB.seq.*
8: /cgcn2_6/ptodata/1/ina/RE.COMB.seq.*
9: /cgcn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	603	100.0	950	3	US-08-098-327E-32 Sequence 32, App1
2	603	100.0	950	3	US-08-462-625-32 Sequence 32, App1
3	603	100.0	1482	3	US-08-098-327E-41 Sequence 41, App1
4	603	100.0	1482	3	US-08-462-625-41 Sequence 41, App1
5	603	100.0	1493	3	US-08-098-327E-38 Sequence 38, App1
6	603	100.0	1493	3	US-08-462-625-38 Sequence 38, App1
7	600	99.5	1482	3	US-08-098-327E-45 Sequence 45, App1
8	600	99.5	1482	3	US-08-462-625-45 Sequence 45, App1
9	564	93.5	1320	2	US-08-257-073-15 Sequence 15, App1

10	111	18.4	71	2	US-08-257-073-130 Sequence 130, App
11	97	16.1	2130	3	US-09-248-796A-4942 Sequence 4942, Ap
12	93	15.4	73519	3	US-09-949-016-16344 Sequence 16344, A
13	93	15.4	105919	3	US-09-949-016-11769 Sequence 11769, A
14	90.5	15.0	4599	2	US-08-431-080-27 Sequence 27, App1
15	90.5	15.0	4599	2	US-08-938-534-27 Sequence 27, App1
16	90.5	15.0	4599	2	US-09-345-294-27 Sequence 27, App1
17	89.5	14.8	696	3	US-09-461-697-193 Sequence 193, App
18	89.5	14.8	699	3	US-09-461-697-189 Sequence 189, App
19	89.5	14.8	717	3	US-09-461-697-187 Sequence 187, App
20	89.5	14.8	774	3	US-09-461-697-185 Sequence 185, App
21	89.5	14.8	819	3	US-09-461-697-184 Sequence 184, App
22	89.5	14.8	1669	3	US-09-461-697-184 Sequence 2899, Ap
23	86	14.3	1344	3	US-09-710-279-2899 Sequence 3225, Ap
24	86	14.3	1353	3	US-09-710-279-3813 Sequence 3813, Ap
25	86	14.3	3001	3	US-09-710-279-3813 Sequence 4245, Ap
26	86	14.3	3974	3	US-09-710-279-4245 Sequence 1, App1
27	85.5	14.2	3739	3	US-09-882-274-1 Sequence 12506, A
28	85	14.1	19728	3	US-09-949-016-12506 Sequence 15767, A
29	85	14.1	19728	3	US-09-949-016-15767 Sequence 2716, Ap
30	84.5	14.0	1785	3	US-09-949-016-2716 Sequence 756, App
31	84.5	14.0	1758	3	US-09-949-016-756 Sequence 1, App1
32	84.5	14.0	3645	2	US-08-663-112-1 Sequence 2, App1
33	84	13.9	5874	3	US-09-561-709B-2 Sequence 1957, Ap
34	83.5	13.8	1685	3	US-09-949-016-16344 Sequence 16344, A
35	83.5	13.8	73519	3	US-09-949-016-11769 Sequence 11769, A
36	83.5	13.8	105919	3	US-09-949-016-11769 Sequence 1, App1
37	82.5	13.7	2471	3	US-09-915-060A-3 Sequence 12498, A
38	81.5	13.5	5788	3	US-09-949-016-12498 Sequence 14458, A
39	81.5	13.5	5788	3	US-09-949-016-14458 Sequence 15247, A
40	81.5	13.5	5870	3	US-10-164-595-53 Sequence 53, App1
41	81	13.4	4248	3	US-09-949-016-13217 Sequence 13217, A
42	81	13.4	28806	3	US-09-949-016-13217 Sequence 1630, App
43	80.5	13.3	431	3	US-09-513-999C-1630 Sequence 521, App
44	80.5	13.3	1353	3	US-08-956-171E-521 Sequence 521, App
45	80.5	13.3	1353	3	US-08-781-986A-521

ALIGNMENTS

RESULT 1
US-08-098-327E-32
Sequence 32, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: DRUHLER, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
City: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300

Db 670 AGAAAAAGAACTGAGATATATATGACAGAGATTATATGCTGTTAGAAATACCA 729
Qy 41 AAlaIleGlueuProSerGIuaenGIuaRgIyTYrTYrIleProHISgInSerSeriou 60
Db 730 GCTTAAAGACTTCATCAGAAATGAAAGCTGATATTATATACCACTCAATCTTCTTTA 789
Qy 61 ProGInaSPaenARgIyaenSerARgAspSerIySgIuIleSeriIleIleGIuLyThr 80
Db 790 CCTCAGGACAAACAGAGGAAATAGTAGATTCCAAGAAATATCTATATAGAAAAACA 849
Qy 81 AsnARgIuSerIleTHrThAsnValGIuGIyARgARgAspIleHISlySGIyHISleu 100
Db 850 AATAGAGATCTATTATCAACAAATGTTGAGAGACGAAAGGATATACATAAAGACATCTT 909
Qy 101 GIuGIuLySlySAspGIySerIleIySProGIuInLySGIuAspLySser 117
Db 910 GAAGAAAAGAAAGATGTTCAATTAACCAAGAAACAAAAGAAATTAATCT 960

RESULT 5
US-08-098-327E-38
Sequence 38, Application US/08098327E
Patent No. 6270771

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098.327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-38

Alignment Scores:

Pred. No.: 9,23e-66 Length: 1493
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-08-098-327E-38 (1-1493)

Qy 1 LeuGIuGIuGIuInnARgAspLeuGIuInARgIyAlaSPThrIySlySAsnLeuGIu 20
Db 607 TTGCAAGACGACGAAAGAGATTATAGAAACAAGAAAGGCTGATACGAAAAAATTTTGAA 666
Qy 21 ArgIySlySgIuHISgIyAspIleuAlaGIuAspLeuYrGIyARgLeuGIuIlePro 40
Db 667 AGAAAAAGAAACATGAGATATATAGACAGAGATTATATGCTGTTAGAAATACCA 726
Qy 41 AAlaIleGlueuProSerGIuaenGIuaRgIyTYrTYrIleProHISgInSerSeriou 60
Db 727 GCTATAGAACTTCATCAAGAAATGAAAGCTGATATTATATACCACTCAATCTTCTTTA 786
Qy 61 ProGInaSPaenARgIyaenSerARgAspSerIySgIuIleSeriIleIleGIuLyThr 80
Db 787 CCTCAGGACAAACAGAGGAAATAGTAGATTCCAAGAAATATCTATATAGAAAAACA 846
Qy 81 AsnARgIuSerIleTHrThAsnValGIuGIyARgARgAspIleHISlySGIyHISleu 100
Db 847 AATAGAGATCTATTATCAACAAATGTTGAGAGACGAAAGGATATACATAAAGACATCTT 906
Qy 101 GIuGIuLySlySAspGIySerIleIySProGIuInLySGIuAspLySser 117
Db 907 GAAGAAAAGAAAGATGTTCAATTAACCAAGAAACAAAAGAAATTAATCT 957

RESULT 6

US-08-462-625-38
Sequence 38, Application US/08462625
Patent No. 6319502

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462.625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098.327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid

RESULT 9
US-08-257-073-15

```

OY 2 GlnGlnGlnGlnArgAspIeuGln-----ArgIysAla 13
Db 427 GAAAAACAGCAAGAAATCTTGTAAGAAAAAGCCGCCGACCTTAAGAAACGAGGAAGCT 486
OY 14 AspThrIlybysAsnIeuGlnuArgIyIsGluHfSGIYAspIleLeuAlaGluAspIeu 33
Db 487 GATACGAAAAAAATTTAGAAAAAGAAAGGAACTGAGATGTATTAGCAGAGATTTA 546
OY 34 TyrGlyArgIeuGlnIleProAlaIleGluIeuProSerGluAsnGluArgGlyTYrTYr 53
Db 547 TATGGTCGTTTAAAAATACCACTATAGAACTTCATCAGAAATGAACCGTGATATTAT 600
OY 54 IleProHisGlnSerSerIeuProGlnAspAsnArgGlyAsnSerArgAspSerIeu 73

```


Db 607 ATACACATCATCTTCTTACTCAGACACAGAGGAATGATGATTCACAAGAA 666
Qy 74 |||ser|||le|||eg|||u|||y|||th|||r|||a|||n|||a|||r|||g|||l|||u|||e|||t|||h|||r|||a|||n|||v|||a|||l|||g|||u|||g|||y|||a|||r|||g|||a|||r|||g||| 93
Db 667 ATATCTATATAGAAAAAACAATAGAGAACTTATCAACAATGTTGAAGGACGAAAG 726
Qy 94 Appl|||e|||t|||e|||r|||g|||y|||h|||e|||s|||e|||u|||g|||u|||l|||y|||a|||b|||s|||e|||r|||e|||y|||s|||p|||r|||o|||g|||u|||n|||t|||y|||s||| 113
Db 727 GATATACATTAAGACACTTGTGAAGAAAGAAAGTGTTCATTAACCAAGAAACAAAAA 786
Qy 114 GluAsp|||y|||s|||e|||r||| 117
Db 787 GAAGATTAATCT 798
RESULT 10
US-08-257-073-130
; Sequence 130, Application US/08257073
; Patent No. 5766597
; GENERAL INFORMATION:
; APPLICANT: Paolucci, Enzo
; APPLICANT: de Taisne, Charles
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th Floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2570
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TEXT: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 130:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-257-073-130
Alignment Scores:
Pred. No.: 4.87e-06 Length: 71
Score: 111.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.41% Indels: 0
Gaps: 0
Db: 2 Gaps: 0

US-09-837-344-20 (1-117) x US-08-257-073-130 (1-71)
Qy 69 ArgAsp|||Ser|||e|||g|||u|||l|||e|||Ser|||e|||l|||e|||g|||u|||y|||t|||h|||r|||a|||n|||a|||r|||g|||l|||u|||e|||t|||h|||r|||a|||n||| 88
Db 1 AGAGATTCACAAAGAAATATCTATTAATGAAAAAACAATAGAAATCTATTACAAACAAT 60
Qy 89 Val|||Glu|||Gly 91
Db 61 GTTGAAAGGA 69
RESULT 11
US-09-248-796A-4942
; Sequence 4942, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4942
; LENGTH: 2130
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-4942

Alignment Scores:

Pred. No.: 0.0294 Length: 2130
Score: 97.00 Matches: 37
Percent Similarity: 44.00% Conservative: 29
Best Local Similarity: 24.67% Mismatches: 36
Query Match: 16.09% Indels: 48
Gaps: 9

US-09-837-344-20 (1-117) x US-09-248-796A-4942 (1-2130)

Qy 11 Arg|||y|||s|||a|||l|||a|||s|||p|||h|||r|||y|||s|||y|||s|||a|||n|||e|||u|||g|||u|||a|||r|||g|||y|||e|||g|||u|||h|||e|||g|||y|||a|||e|||p|||l|||e|||----- 28
Db 1384 AAGAGATGATAGTACTGATTTCTTAATATTCAGAAAGAGAAATTTATTCCTCAATTAAGAG 1443
Qy 29 -----Leu|||A|||g|||u|||a|||s|||p|||e|||u|||r|||g|||y|||a|||r|||g|||e|||u|||g|||u|||l|||e|||p|||r|||o|||a||| 41
Db 1444 CCACTAAAGATGAAAAACAATTAAGAAAGACTTGCTC---AAAGTTGAACCCATCC 1500
Qy 42 -----|||e|||g|||u|||l|||e|||u|||p|||r|||o|||s|||e|||r|||g|||u|||a|||n|||g|||u|||a|||r|||g|||y|||r|||y|||r|||i|||e||| 54
Db 1501 AAGATGAGACCGCGGAGGTTGAAGTTCATCCAAAGATGAACACCATTCACAGATT 1560
Qy 55 Pro|||h|||e|||g|||i|||n|||s|||e|||r|||e|||u|||p|||r|||o|||g|||i|||n|||a|||s|||p|||a|||r|||g|||y|||a|||n|||s|||e|||r|||g|||a|||s|||p|||e|||r|||y|||s|||----- 72
Db 1561 CAAATATTTGATAGT---ACCACGTGATAT---GCAAACTTAAGATGAGAAAGATTAAT 1614
Qy 73 -----Glu|||l|||e|||s|||e|||r|||e|||l|||e|||g|||u|||y|||t|||h|||r|||a|||n|||a|||r|||g|||u|||s|||e|||r||| 84
Db 1615 GTTCTTAAGTGTATTTCAAGTTGATACATTAAGTATGTTACCAAGACATCAAGATGAT 1674
Qy 85 Ile-----Th|||r|||h|||a|||n|||v|||a|||l|||g|||u|||g|||y|||a|||r|||g|||y|||a|||s|||p||| 94
Db 1675 ATTAAAAAGTAGAAGAACCTGTGAAAGAAAGACGACGACGAAAGGTAAAGAAAGAC 1734
Qy 95 -----|||h|||e|||l|||y|||e|||g|||y|||h|||e|||s|||e|||u|||g|||u|||l|||y|||a|||b|||s|||e|||r|||e|||y|||s|||p|||r|||o|||g|||u|||n|||t|||y|||s||| 109
Db 1735 AATGAAAAAGTTACTGTTCAACAGACAAATGAGAGAAATGATTAAGATGCGTCAAT 1794
Qy 110 Pro|||Glu|||-----G|||n|||t|||y|||e|||g|||u|||a|||s|||p||| 115
Db 1795 GAAGCCAAAGATAGCTTCAAGAAAGATGAT 1824

```

RESULT 12
US-09-949-016-16344
; Sequence 16344, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16344
; LENGTH: 73519
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(73519)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16344

```

Alignment Scores:	
Pred. No.:	12
Score:	93.00
Percent Similarity:	45.22%
Best Local Similarity:	29.57%
Query Match:	15.42%
DB:	3
Length:	73519
Matches:	34
Conservative:	18
Mismatches:	49
Indels:	14
Gaps:	4

US-09-837-344-20 (1-117) x US-09-949-016-16344 (1-73519)

[illegible]

RESULT 13
 US-09-949-016-11769
 : Sequence 11769, Application US/09949016
 : Patent No 6812338
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: VENTER, J. Craig et al.
 :
 : TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 : TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 : FILE REFERENCE: CL001307

```

1 CURRENT APPLICATION NUMBER: 05/09/949,016
2
3 CURRENT FILING DATE: 2000-04-14
4
5 PRIOR APPLICATION NUMBER: 60/241,755
6
7 PRIOR FILING DATE: 2000-10-20
8
9 PRIOR APPLICATION NUMBER: 60/237,768
10
11 PRIOR FILING DATE: 2000-10-03
12
13 PRIOR APPLICATION NUMBER: 60/231,498
14
15 PRIOR FILING DATE: 2000-09-08
16
17 NUMBER OF SEQ ID NOS: 207012
18
19 SOFTWARE: FASTSQ for Windows Version 4.0
20
21 SEQ ID NO 11769
22
23 LENGTH: 105919
24
25 TYPE: DNA
26
27 ORGANISM: Human
28
29 FEATURE:
30 NAME/KEY: misc_feature
31 LOCATION: (1)...(105919)
32
33 OTHER INFORMATION: n = A,T,C or G
34
35 US-09-949-016-11769

```

Alignment Scores:	
Pred. No.:	19, 9
Score:	93, 00
Percent Similarity:	45, 22%
Best Local Similarity:	29, 57%
Query Match:	15, 42%
DB:	3
Gaps:	4
Length:	10591g
Matches:	34
Conservative:	18
Mismatches:	49
Indels:	14
Gaps:	4

US-09-837-344-20 (1-117) x US-09-949-016-11769 (1-105919

[illegible]

```

1  RESULT 14
2  /
3  / Sequence 27, Application US/06431080
4  / Patent No. 569868
5  /
6  / GENERAL INFORMATION:
7  /
8  / APPLICANT: Gottschling, Daniel E.
9  /
10 / APPLICANT: Singer, Miriam S.
11 /
12 / TITLE OF INVENTION: Telomerase Compositions and Methods
13 /
14 / NUMBER OF SEQUENCES: 32
15 /
16 / CORRESPONDENCE ADDRESS:
17 /
18 / ADDRESSEE: Arnold, White & Durkee
19 /
20 / STREET: P.O. Box 4433
21 /
22 / CITY: Houston
23 /
24 / STATE: TEXAS
25 /
26 / COUNTRY: UNITED STATES OF AMERICA
27 /
28 / ZIP: 77210
29 /
30 / COMPUTER READABLE FORM:
31 /
32 / MEDIUM TYPE: Floppy disk
33 /
34 / COMPUTER: IBM PC compatible
35 /

```

```

OPERATING SYSTEM: PC-DOS/MC-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,080
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Parker David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ACD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4599 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-431-080-27

Alignment Scores:
Pred. No.: 0.55 Length: 4599
Score: 90.50 Matches: 31
Percent Similarity: 34.90% Conservative: 21
Best Local Similarity: 20.81% Mismatch: 54
Query Match: 15.01% Indels: 43
DB: Gaps: 4

US-09-837-344-20 (1-117) x US-08-431-080-27 (1-4599)
Qy 3 GIUGINGIARGAPBLEUGLGNALRGLYSLAEPHTRLYBLVABNLEUGIARGLYS 22
Db 3427 GATAGATGAAAGAAGCTCAACACGAAAAAACTGAAAAAAAAAGCGTTAAAGAAAAA 3486
Qy 23 LYSGLU----- 24
Db 3487 AAGAAGCTTCTTAATAAGAAAACAAGCAAAGCAATGAAGAGCAAGAAACTATG 3546
Qy 25 -----HISGLYASP 27
Db 3547 AATTACCAATTGGGAATCAATGCCCATGATCATCCGTAAACAATTAACGCCTATGCAC 3606
Qy 28 ILE--LEUALGLAPBLEUTRYGLYRGLIENGULILEPROIALLEGLEUPROSER 46
Db 3607 ATAAATCACCGATCCGATTTTACACCAATGAAMATAACCCTATGAATGAATTCCTCT 3666
Qy 47 GLU---AANGIARGGLYTRYTYRIIEPROHISGLINSERSEULEUPROGLNASPARNRG 65
Db 3667 CACGCACTTGAAAGATCCGTCATTAATTAACCTCATTAATCTGATCTTGCCGTGACACGAT 3726
Qy 66 GLAENSERARGAPSPERYSGLIUE-----SERILEGLUYYS 79
Db 3727 ACAAGAAAAATTCAACAAAAGTGTGGTTTAGTAGTAATTCATGAATTTTGGCAA 3786
Qy 80 THRANAARGIUSERILETHRTHRANVALGIUGIYARGATGARPILLEILYSGLYHS 99
Db 3787 GATGAAAAATGACTTACTGCTCTGTAGTAGTATTAACGGTTATGATGACAAAGAGTAT 3846
Qy 100 LEUGIUGLYSLYBAPGLYSERILE 108
Db 3847 GTGATCGAAGATATGACGCGCATATC 3873

RESULT 15
US-08-938-534-27
Sequence 27, Application US/08938534
Patent No. 5916752
GENERAL INFORMATION:
```

```

1  APPLICANT:  Gotschling, Daniel E.
2  TITLE OF INVENTION:  Telomerase Compositions and Methods
3  NUMBER OF SEQUENCES:  32
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE:  Arnold, White & Durkee
6  STREET:  P.O. Box 4433
7  CITY:  Houston
8  STATE:  TEXAS
9  COUNTRY:  UNITED STATES OF AMERICA
10 ZIP:  77210
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE:  Floppy disk
14 COMPUTER:  IBM PC compatible
15 OPERATING SYSTEM:  PC-DOS/MS-DOS/ASCII
16 SOFTWARE:  PatentIn Release #1.0, Version #1.30
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER:  US/08/938,534
19 FILING DATE:  26-SEP-1997
20 CLASSIFICATION:  536
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER:  08/431,080
23 FILING DATE:
24 APPLICATION NUMBER:  SN 08/326,781
25 FILING DATE:  October 20, 1994
26 ATTORNEY/AGENT INFORMATION:
27 NAME:  Parker, David L.
28 REGISTRATION NUMBER:  32,165
29 REFERENCE/DOCKET NUMBER:  ARCD:155/PAR
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE:  (512) 418-3000
32 TELEFAX:  (713) 789-2679
33
34 INFORMATION FOR SEQ ID NO:  27:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH:  4599 base pairs
37 TYPE:  nucleic acid
38 STRANDEDNESS:  single
39 TOPOLOGY:  linear
40
41 US-08-938-534-27
42
43 Alignment Scores:
44 Pred. No.:  0.55          Length:  4599
45 Score:  90.50          Matches:  31
46 Percent Similarity:  34.90%      Conservative:  21
47 Best Local Similarity:  20.81%   Mismatches:  54
48 Query Match:  15.01%          Indels:  43
49 DBs:  2          Gaps:  4
50
51 US-09-837-344-20 (1-117) x US-08-938-534-27 (1-4599)
52
53 3  GUGGNGlAaRgaAPLeGluGluARgLySa1aSPThrLySLySaEnLeuGluARgLyS 22
54 3427  GATTAAGATGAAGAAGACTTCAACGCAAAAAAACTGAAAAAAAGGCACTTAAAGAAAAA 3486
55
56 23  LysGlu----- 24
57 3487  AAGAAGCTTCTTAAATAATGAAGAACAAGCAAAAGGCATPAAGAGCAAGAATCATG 3546
58
59 25 -----HisGlyASP 27
60 3547  AATTTCACATTGGGATCATGCGCATGAGTCATCGGTACAAATAACAGCCATAGCGAC 3606
61
62 28  Ile---LeuAlaGluAPLeuTyrgLYaRgLeuGluIleProAlaIleGluLeuProSer 46
63 3607  ATTAATACCGGTACCGGATTTTTCACAACCAATGAATACCCCTTGAATGAATGAACTTCCCTCT 3666
64
65 47  Glu---AenGluARgLYTyTrrIleProHisGlnSerSerLeuProGlnASPasnARg 65
66 3667  CACGCACTGAAGAATGCGCTCATTAATACCTCATATTCTGATCTTGCCGTGAGACAGCAAT 3726
67
68 66  GlyAenSerARgASPserLyGluIle-----SerIleIleGluIle 79

```

```

Db      3727 ACAAGGAAAAATTCACAAAAAGCTGTGTTAGATGAATTCAATGAGATTTGGGCAA 3786
Qy      80 ThrAsnArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHisGlyHis 99
          ::: ::::
Db      3787 GATGAAAAATGACTTACTGTCTGTAGTGATATTAAAGGTTATGATGACAAAGAGGTGAT 3846
          GATGAAAAATGACTTACTGTCTGTAGTGATATTAAAGGTTATGATGACAAAGAGGTGAT 3846
Qy      100 LeuGluGluLysLysAspGlySerIle 108
          ::: ||| |||
Db      3847 GTGATCGAAGATATCTGACCGCGATATC 3873
          GTGATCGAAGATATCTGACCGCGATATC 3873

```

Search completed: December 7, 2005, 12:45:14
 Job time : 127.951 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 10:33:33 ; Search time 409.905 Seconds
(without alignments)
2360.343 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603

Sequence: 1 LQEQRDLEGRKADTKKLE.....GHLEKKDGSIKPEQKEDKS 117

Scoring table: BL0SUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.spool -DEV=xlh
-O/cgn2_1/USPTO.spool h/US09837344/runat 05122005 133535 839/app query.fasta_1.718
-DB=Published_Applications_NA_Main -QMT=Fastcap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=dl0sum62
-TRANS=numa40.cdi -LIST=45 -DCCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09837344 @CGN 1.1 1364 @runat 05122005 133535_839
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOC -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAP=10 -XGAPEXT=0.5
-FGAP=6 -FGAPEXT=7 -YGAP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications_NA_Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	603	100.0	950	US-09-837-344-32	Sequence 32, Appl
2	603	100.0	950	US-09-900-963-32	Sequence 32, Appl
3	603	100.0	1482	US-09-837-344-41	Sequence 41, Appl
4	603	100.0	1482	US-09-900-963-41	Sequence 41, Appl
5	603	100.0	1493	US-09-837-344-38	Sequence 38, Appl
6	603	100.0	1493	US-09-900-963-38	Sequence 38, Appl
7	600	99.5	1482	US-09-837-344-45	Sequence 45, Appl
8	600	99.5	1482	US-09-900-963-45	Sequence 45, Appl

9	94	15.6	1385	3	US-09-894-018-122	Sequence 122, App
10	94	15.6	1385	8	US-10-474-960A-122	Sequence 122, App
11	90.5	15.0	969	9	US-10-450-763-7484	Sequence 7484, Ap
12	90.5	15.0	969	9	US-10-450-763-11162	Sequence 11162, A
13	90.5	15.0	969	9	US-10-450-763-28063	Sequence 28063, A
14	90	14.9	4080	6	US-10-172-118-1520	Sequence 1520, Ap
15	90	14.9	4080	7	US-10-342-887-1520	Sequence 1520, Ap
16	90	14.9	4080	9	US-10-756-149-1669	Sequence 1669, Ap
17	89.5	14.8	666	3	US-09-922-261-193	Sequence 193, App
18	89.5	14.8	699	3	US-09-922-261-191	Sequence 191, App
19	89.5	14.8	717	3	US-09-922-261-189	Sequence 189, App
20	89.5	14.8	774	3	US-09-922-261-187	Sequence 187, App
21	89.5	14.8	793	3	US-09-922-261-185	Sequence 185, App
22	89.5	14.8	1669	3	US-09-922-261-184	Sequence 184, App
23	87.5	14.5	700	5	US-10-040-739-1385	Sequence 1385, Ap
24	87	14.4	903	6	US-10-369-493-29682	Sequence 29682, A
25	86.5	14.3	601	7	US-10-021-323-563	Sequence 563, App
26	86.5	14.3	3684	10	US-11-097-143-32615	Sequence 32615, A
27	85.5	14.2	2298	7	US-10-052-482-186	Sequence 186, App
28	85.5	14.2	3477	3	US-09-969-347-221	Sequence 221, App
29	85.5	14.2	3477	9	US-10-843-641A-8350	Sequence 8350, Ap
30	85.5	14.2	3739	3	US-09-882-274-1	Sequence 1, Appl1
31	85.5	14.2	3740	7	US-10-052-482-185	Sequence 185, App
32	85.5	14.2	3740	8	US-10-484-577-680	Sequence 680, App
33	85.5	14.2	3756	6	US-10-133-013-255	Sequence 255, App
34	85	14.1	613	7	US-10-424-599-28673	Sequence 28673, A
35	85	14.1	11065	8	US-10-723-860-2002	Sequence 2002, Ap
36	85	14.1	11065	9	US-10-756-149-1888	Sequence 1888, Ap
37	84.5	14.0	729	9	US-10-450-763-2486	Sequence 2486, Ap
38	84.5	14.0	1758	9	US-10-450-763-11623	Sequence 11623, A
39	84.5	14.0	2108	3	US-09-962-832-225	Sequence 225, App
40	84.5	14.0	2108	3	US-10-240-425-1294	Sequence 1294, App
41	84.5	14.0	2108	9	US-10-843-641A-6111	Sequence 6111, Ap
42	84.5	14.0	3645	3	US-09-292-758-85	Sequence 85, Appl
43	84.5	14.0	3645	6	US-10-388-410-4	Sequence 4, Appl1
44	84.5	14.0	630	4	US-09-925-065A-788059	Sequence 788059,

ALIGNMENTS

RESULT 1
US-09-837-344-32
Sequence 32, Application US/09837344
Patent No. US2002004182A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUTHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-837-344-32

Alignment Scores:
Pred. No.: 4,27e-66 Length: 950
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-09-837-344-32 (1-950)

QY 1 LeuGInGInGInGInaRgAspLeuGInaRgLyAlaAspThrLyLyAsnLeuGIn 20
DB 600 TTGCAAGACGACGACAAAGAGATTAGAACAAAGAGAGCTGATACGAAAAAATTTAGAA 659

QY 21 ArgLyLyGInGInGInGInaRgAspLeuGInaRgLyAlaAspThrLyLyArgLeuGInLeuPro 40
DB 660 AGAAAAAGAGACATGAGATATATTAGCAGAGATTATATGTCGTTTGAATAACCA 719

QY 41 AlaIleGInLeuProSerGInaRgLyAlaAspThrLyLyLeuProHisGInSerSerLeu 60
DB 720 GCTATGAACTTCATCAGAAATGAACTGATATTATATACCATCATCTCTTTA 779

QY 61 ProGInaRgAspAsnRgLyAsnSerArgAspSerLyGInLeuSerIleIleGInLyThr 80
DB 780 CCTCAGACACACAGAGGAGATAGATTCACAGAAATATCTTATATAGAAAAACA 839

QY 81 AsnArgGInSerIleThrThraenValGInGInaRgAspIleHisLyGInHisLeu 100
DB 840 AATAGAGATCTATTCAACAAATGTTGAAGACGAGGATATACATTAAGACATCTT 899

QY 101 GInGInLyLyAspGInSerIleLyProGInGInLyGInaRgLySer 117
DB 900 GAAGAAAGAGAGATGTTCAATMAAACGAAACAAAGAGATMAATCT 950

RESULT 2
US-09-900-963-32
Sequence 32, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burne, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-900-963-32

Alignment Scores:
Pred. No.: 4,27e-66 Length: 950
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-09-900-963-32 (1-950)

QY 1 LeuGInGInGInGInaRgAspLeuGInaRgLyAlaAspThrLyLyAsnLeuGIn 20
DB 600 TTGCAAGACGACGACAAAGAGATTAGAACAAAGAGAGCTGATACGAAAAAATTTAGAA 659

QY 21 ArgLyLyGInGInGInGInaRgAspLeuGInaRgLyAlaAspThrLyLyArgLeuGInLeuPro 40
DB 660 AGAAAAAGAGACATGAGATATATTAGCAGAGATTATATGTCGTTTGAATAACCA 719

QY 41 AlaIleGInLeuProSerGInaRgLyAlaAspThrLyLyLeuProHisGInSerSerLeu 60
DB 720 GCTATGAACTTCATCAGAAATGAACTGATATTATATACCATCATCTCTTTA 779

QY 61 ProGInaRgAspAsnRgLyAsnSerArgAspSerLyGInLeuSerIleIleGInLyThr 80
DB 780 CCTCAGACACACAGAGGAGATAGATTCACAGAAATATCTTATATAGAAAAACA 839

QY 81 AsnArgGInSerIleThrThraenValGInGInaRgAspIleHisLyGInHisLeu 100
DB 840 AATAGAGATCTATTCAACAAATGTTGAAGACGAGGATATACATTAAGACATCTT 899

QY 101 GInGInLyLyAspGInSerIleLyProGInGInLyGInaRgLySer 117
DB 900 GAAGAAAGAGAGATGTTCAATMAAACGAAACAAAGAGATMAATCT 950

RESULT 3
US-09-837-344-41
Sequence 41, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE

Db 610 TTGCAAGACGACAAAGATTAGAACAAAGAAAGCTGATACGAAAAAATTAGAA 669
Qy 21 ArglyblysglunhiglyaspdlleleuagluaspdeuTyrglyArgleugluilepro 40
Db 670 AAAAAAAGAAACATGAGATATATTAGACAGAGATTATATGCTGTTAAATAACCA 729
Qy 41 AlaileguleuProsergluansgluArglyTyrylleProhieglnserleu 60
Db 730 GCTATGAACTTCATCAGAAATGACCTGATATATATACCATCAATCTCTTTA 789
Qy 61 Proglinaapanaargglyanserargaspserlysgluileserlleleuglysthr 80
Db 790 CCTCAGGACAAACAGAGGAGATAGTAGATTCAGAGAAATATCTATATAAGAAAAACA 849
Qy 81 AanaarglyserllethThrasnValgluglyArgargaspdllehislysglyhisleu 100
Db 850 AATAGGAATCTATTACAACAATGTTGAAGACGAAAGGATATATATTAAGACATCTT 909
Qy 101 GluglyblysglyaspglyserllelyProgluglnlygluaspbyser 117
Db 910 GAAGAAAAAGAAAGATGCTTCATTAACCAAGAAACAAAAAGATTAATCT 960

RESULT 5
US-09-837-344-38
; Sequence 38, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1493 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 38:
; US-09-837-344-38

Alignment Scores:
Pred. No.: 7,686-66 Length: 1493
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-09-837-344-38 (1-1493)

Qy 1 Leuglnluginlgnaraspdeugluinarlysaiaapthlyblyasnleuglu 20
Db 607 TTGCAAGACGACAAAGATTAGAACAAAGAAAGCTGATACGAAAAAATTAGAA 666
Qy 21 ArglyblysglunhiglyaspdlleleuagluaspdeuTyrglyArgleugluilepro 40
Db 667 AAAAAAAGAAACATGAGATATATTAGACAGAGATTATATGCTGTTAGAAATACCA 726
Qy 41 AlaileguleuProsergluansgluArglyTyrylleProhieglnserleu 60
Db 727 GCTATGAACTTCATCAGAAATGACCTGATATATATACCATCAATCTCTTTA 786
Qy 61 Proglinaapanaargglyanserargaspserlysgluileserlleleuglysthr 80
Db 787 CCTCAGGACAAACAGAGGAGATAGTAGATTCAGAGAAATATCTATATAGAAAAACA 846
Qy 81 AanaarglyserllethThrasnValgluglyArgargaspdllehislysglyhisleu 100
Db 847 AATAGGAATCTATTACAACAATGTTGAAGACGAAAGGATATATTAAGACATCTT 906
Qy 101 GluglyblysglyaspglyserllelyProgluglnlygluaspbyser 117
Db 907 GAAGAAAAAGAAAGATGCTTCATTAACCAAGAAACAAAAAGATTAATCT 957

RESULT 6
US-09-900-963-38
; Sequence 38, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,963
; FILING DATE: 10-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/098,327
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:

LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-900-963-38

Alignment Scores:

Pred. No.:	7,686-66	Length:	1493
Score:	603.00	Matches:	117
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-09-837-344-20 (1-117) x US-09-900-963-38 (1-1493)

QY 1 LeuGluGluGlnGlnAArgApLeuGluGlnAArgAAlaAspThrIlySlySaenLeuGlu 20
DB 607 TTCCAAAGAGCAGCAAGAGATTGAGACAAAGAGAGCTGATACGAAAAAATTTTGA 666
QY 21 ArgLylySgIuhIsgIyAspIleuAlaGluAspLeuTyrgIyArgLeuGluIlePro 40
DB 667 AGAAAAAGCAACATGAGATATATTAGCAGAGATTATATGCTGTTAGAAATACCA 726
QY 41 AlaIleGluLeuProSerGluAenGluArgIyTyrrIleProHsgInserSerIeu 60
DB 727 GCTATAGAACTTCCATCAGAAAAAGAACGTGATATATATCCACATCATCTTCTT 786
QY 61 ProGlnAspAenArgIyAenSerArgAspSerIySgIleSerIleIleGluIySthr 80
DB 787 CCTCAGACAAACAGAGGAATAGTAGATTCCAGGAATATCTTAATATGAAAAACA 846
QY 81 AsnArgIuSerIleThrAsnValGluGlyArgAspIleHslySgIyHsleu 100
DB 847 AATAGAGATCTTATACAAACAACTTGGAAGAGAGGATATACATTAAGACATCTT 906
QY 101 GluGluIylySAspGlySerIleLysProGluGlnIySgIuAspLysSer 117
DB 907 GAAGAAAAAGAAAGATGTTCAATATAAACCAAGAACAAAGAAATTAATCT 957

RESULT 7

US-09-837-344-45
Sequence 45, Application US/09837344
Patent No. US20020041882A1

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine

DRUILHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE

HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE

OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathie

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/837,344

FILING DATE: 19-Apr-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CDS

LOCATION: 1..1482

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992

SEQUENCE DESCRIPTION: SEQ ID NO: 45:

US-09-837-344-45

US-09-837-344-45

Pred. No.:	1,826-65	Length:	1482
Score:	600.00	Matches:	116
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.15%	Mismatches:	0
Query Match:	99.50%	Indels:	0
DB:	3	Gaps:	0

US-09-837-344-20 (1-117) x US-09-837-344-45 (1-1482)

QY 1 LeuGluGluGlnGlnAArgApLeuGluGlnAArgAAlaAspThrIlySlySaenLeuGlu 20
DB 610 TTCCAAAGAGCAGCAAGAGATTGAGACAAAGAGAGCTGATACGAAAAAATTTGAA 669
QY 21 ArgLylySgIuhIsgIyAspIleuAlaGluAspLeuTyrgIyArgLeuGluIlePro 40
DB 670 AGAAAAAGCAACATGAGATATATTAGCAGAGATTATATGCTGTTAGAAATACCA 729
QY 41 AlaIleGluLeuProSerGluAenGluArgIyTyrrIleProHsgInserSerIeu 60
DB 730 GCTATAGAACTTCCATCAGAAAAAGAACGTGATATATATCCACATCATCTTCTT 789
QY 61 ProGlnAspAenArgIyAenSerArgAspSerIySgIleSerIleIleGluIySthr 80
DB 790 CCTCAGACAAACAGAGGAATAGTAGATTCCAGGAATATCTTAATATGAAAAACA 849
QY 81 AsnArgIuSerIleThrAsnValGluGlyArgAspIleHslySgIyHsleu 100
DB 850 AATAGAGATCTTATACAAACAACTTGGAAGAGAGGATATACATTAAGACATCTT 909
QY 101 GluGluIylySAspGlySerIleLysProGluGlnIySgIuAspLysSer 117
DB 910 GAAGAAAAAGAAAGATGTTCAATATAAACCAAGAACAAAGAAATTAATCT 960

RESULT 8

US-09-900-963-45
Sequence 45, Application US/09900963
Publication No. US20030064075A1

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine

DRUILHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE

HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE

OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46


```

; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1385
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encoding sequence for PECTL/HTL(N)
US-10-474-960A-122

Alignment Scores:
Pred. No.: 0.161 Length: 1385
Score: 94.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 15.59% Indels: 0
DB: Gaps: 0

US-09-837-344-20 (1-117) x US-10-474-960A-122 (1-1385)

Qy 40 ProballeleugluLeuProsergluAenGluArgGlyTyrIleProHisGlnSerSer 59
Db 361 CTCGACCAAGCTCTCGCGTCCGAAACGAAAGAGATTAATCACTCCACCAAGACAGC 420

Qy 60 Leu 60
Db 421 CTC 423

RESULT 11
US-10-450-763-7484/c
; Sequence 7484, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 7484
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (482)..(183)
; OTHER INFORMATION: 51% homologous to Homo sapiens p40, accession number
; OTHER INFORMATION: U93569, Smith-Waterman Score=226.
US-10-450-763-7484

Alignment Scores:
Pred. No.: 0.28 Length: 969
Score: 90.50 Matches: 30
Percent Similarity: 47.66% Conservative: 21
Best Local Similarity: 28.04% Mismatches: 35
Query Match: 15.01% Indels: 21
DB: Gaps: 6

US-09-837-344-20 (1-117) x US-10-450-763-7484 (1-969)

Qy 1 LeuGlnGluGlnGlnArgAAspleuGluGlnArgGlySerIleProHisGlnLeuGlu 20
Db 600 CTCACGACGCTCCACAGACCTGACGCTGAGGCTCTGACTGTTAAGAGAAAATAAC 541
```

```

Qy 21 ArgGlyLeuGluHisGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGluIlePro 40
Db 540 AAACAGAAAGAGATCCACAG-----CAAAACCC 511

Qy 41 AlaIleGluLeuProSerGluAenGluArgGlyTyrIleProHisGln 57
Db 510 ATCTGTAACGTACCATCATCAAGACCAAGATATTATCCAGAGAACTTCCCAACT 451

Qy 58 SerSerLeuProGln---AspAsnArgGlyAsnSerArgAspSerIleSerIle 76
Db 450 AGCAAGACAGGCCAGCATTCAAATTCCAGGAAATACACAGATGCCCAAGATA---CTC 394

Qy 77 IleGluLeuThrAsnArgIleSerIleThrAsnValGluGlyArgArgAspIleHis 96
Db 393 CTCGAGAAAGGACAC-----TCCAGACACATTAATTGTCCAG-----ATTCCAC 352

Qy 97 LysGlyHisLeuGluGlyLys 103
Db 351 CAAGTTGAATGAAGGAAAAA 331

RESULT 12
US-10-450-763-11162/c
; Sequence 11162, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 11162
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (200)..(430)
; OTHER INFORMATION: 100% homologous to Homo sapiens proteasome subunit
; OTHER INFORMATION: XAPC7, accession number AF022815, Smith-Waterman Score=404.
US-10-450-763-11162

Alignment Scores:
Pred. No.: 0.28 Length: 969
Score: 90.50 Matches: 30
Percent Similarity: 47.66% Conservative: 21
Best Local Similarity: 28.04% Mismatches: 35
Query Match: 15.01% Indels: 21
DB: Gaps: 6

US-09-837-344-20 (1-117) x US-10-450-763-11162 (1-969)

Qy 1 LeuGlnGluGlnGlnArgAAspleuGluGlnArgGlySerIleProHisGlnLeuGlu 20
Db 600 CTCACGACGCTCCACAGACCTGACGCTGAGGCTCTGACTGTTAAGAGAAAATAAC 541

Qy 21 ArgGlyLeuGluHisGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGluIlePro 40
Db 540 AAACAGAAAGAGATCCACAG-----CAAAACCC 511

Qy 41 AlaIleGluLeuProSerGluAenGluArgGlyTyrIleProHisGln 57
Db 510 ATCTGTAACGTACCATCATCAAGACCAAGATATTATCCAGAGAACTTCCCAACT 451

Qy 58 SerSerLeuProGln---AspAsnArgGlyAsnSerArgAspSerIleSerIle 76
```

```
Db 450 ACCAAGACAGCCGACATTCATAATTCAGAAATACAGAAATGCCAAGATA---CTC 394
Qy 77 lTegluYthrAaargluSerileThrAaValgluGlYArgrAaplleHis 96
Db 393 CTCGAGAAAGCAGC-----TCCAGACACATTAATTGTGAG-----ATTGAC 352
Qy 97 lYeglyHleugluGlYs 103
Db 351 CAAGTTGAATGAGAGAAAA 331

RESULT 13
US-10-450-763-28063
; Sequence 28063, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 28063
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (488) ..(787)
; OTHER INFORMATION: 51% homologous to Homo sapiens p40, accession number
; OTHER INFORMATION: U93569, Smith-Waterman Score=226.
US-10-450-763-28063

Alignment Scores:
Pred. No.: 0.28 Length: 969
Score: 90.50 Matches: 30
Percent Similarity: 47.66% Conservative: 21
Best Local Similarity: 28.04% Mismatches: 35
Query Match: 15.01% Indels: 21
DB: 9 Gaps: 6

US-09-837-344-20 (1-117) x US-10-450-763-28063 (1-969)
Qy 1 lAeuglIngluInAargleuGlInAargluAaPThrlyAaenleuGl 20
Db 370 CTCGAGAAAGCTCCAGACAGCTGAGCTGAGCTGTGAGAGAAATTAAC 429
Qy 21 ArglYbelyeGlInleuGlAaPilleuAglAaPleuYrlyArgrleuGlilleP 40
Db 430 AAACAGAAAGACATCCACAG-----CAAAACCC 459
Qy 41 AlAlleuPProsergluAaenGlYArgrlyr-----lIeProHleuGl 57
Db 460 ATCTGTATCTCATTATCAAGAACCAAGATATTATCCAGAGAACTCCCAACCT 519
Qy 58 SerleuPProglu--AaPAAaarglyAaenSerArgrAaPserlyeGlilleSerlle 76
Db 520 AGCAAGACAGCCGACATTCATAATTCCAGAAATACAGAAATGCCAAGATA---CTC 576
Qy 77 lTegluYthrAaargluSerileThrAaValgluGlYArgrAaPilleHis 96
Db 577 CTCGAGAAAGCAGC-----TCCAGACACATTAATTGTGAG-----ATTGAC 618
Qy 97 lYeglyHleugluGlYs 103
Db 619 CAACTGAATGAGAGAAAA 639
```

```
RESULT 14
US-10-172-118-1520
; Sequence 1520, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1520
; LENGTH: 4080
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_014700
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1520

Alignment Scores:
Pred. No.: 2.1 Length: 4080
Score: 90.00 Matches: 31
Percent Similarity: 41.67% Conservative: 24
Best Local Similarity: 23.48% Mismatches: 53
Query Match: 14.93% Indels: 24
DB: 6 Gaps: 4

US-09-837-344-20 (1-117) x US-10-172-118-1520 (1-4080)
Qy 2 gIngluIngluInAargleuGlInAargluAaPThrlyAaenleuGl 21
Db 2167 CAGGAGAACAGAGAGAAATGGGGAGCAGGCTGAGTCAAGAGCACCAGTTCCAGAG 2226
Qy 22 lYslyeGlInleuGlAaPilleuAglAaPleuYrlyArgrleuGlilleP 40
Db 2227 GACAAAGAGCCAGCCAGGAGCTGATCGAGGCTCCGAAAGCAGCTGAGCACTGAG 2286
Qy 41 AlAlleuPProsergluAaenGlYArgrlyr-----lIeProHleuGl 51
Db 2287 CTCCTCAAGCTGAGAGCCAGAGCGCGCGGCGCCAGAGAGCATGGGCTGCAGAG 2346
Qy 52 lYrlyrleProHleuGlAaPserleuPProgluAaP 63
Db 2347 TACCAAGCGCGCGCGGAGAGAGCTGAGCAGAGTCCGAGCTGAGAGCAGAG 2406
Qy 64 AaAarglyAaenSerArgrAaPserlyeGlilleSerlleleuYthrAaarglu 83
Db 2407 AACCCGACACCTGAGAGAGAGAGAGAGCTG-----AACGGGCGAG 2448
Qy 84 SerleuPProglu--AaPAAaarglyAaenSerArgrAaPilleleuYthrAaarglu 103
Db 2449 ATCATTTACCTGAGCATTCAGAGGCGCCAGAGACCTTTCTCCAGAGCTTCTGAGTCC 2508
Qy 104 lYsAaPglYSerlleuPProgluInglYs 115
Db 2509 CTCGCTGAGAGATCAGCTCCGCTCCGAGATGAG 2544

RESULT 15
US-10-342-887-1520
; Sequence 1520, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Lindsey, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIORITY FILING DATE: 2003-01-15
; PRIORITY FILING DATE: 2001-06-18
; PRIORITY FILING DATE: 2001-06-18
; PRIORITY FILING DATE: 2002-05-14
; PRIORITY FILING DATE: 2002-05-14
; PRIORITY FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1520
; LENGTH: 4080
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1520

Alignment Scores:
Pred. No.: 2.1          Length: 4080
Score: 90.00          Matches: 31
Percent Similarity: 41.67%      Conservative: 24
Best Local Similarity: 23.48%    Mismatches: 53
Query Match: 14.93%            Indels: 24
DB: 7                        Gaps: 4

US-09-837-344-20 (1-117) x US-10-342-887-1520 (1-4080)

QY      2  GlnGlnGlnGlnArgAspLeuGlnGlnArgLysAlaSerThrLysLysAsnLeuGluArg 21
      |||||  ::::|  ::::|  |||  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|
DB      2167  CAGAGAGACAGAGAGAGATGGGAGACGCTGAGTCAGAGAGGACCACTTCAGAGG 2226

QY      22  LysLysGlnHisGlyAspIleLeuAlaGlnAspLeuTyrgLysArgLeuGlu---IlePro 40
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ::
DB      2227  GACAAAGAGGCCACCCAGAGCTGATCGAGGACCTCCGAAGCAGCTGGAGCAGCTGCAG 2286

QY      41  AlaIleGlnLeuProSerGlnAsnGlnArgGly----- 51
      ::::|  ::::|  ::::|  |||||  |||||  |||||  |||||  |||||  |||||  ::
DB      2287  CTCCTCAAGCTGAGAGCGGAGCAGCGGCGGCGGCGAGCAGACAGCATGGCCTGCAGAG 2346

QY      52  TyrTyrIleProHisGlnSerSer-----LeuProGlnAsp 63
      |||||  ::::|  ::::|  |||||  |||||  |||||  |||||  |||||  |||||  ::
DB      2347  TACCACAGCCCGCCCGGAGAGCGAGCTGAGCAGAGAGTCCGAGGCTGAAGCAGGAG 2406

QY      64  AsnArgGlyAsnSerArgAspSerLysGlnIleSerIleIleGlnLysThrAsnArgGlu 83
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ::
DB      2407  AACCGCAACTGAAGGAGCAGAAACGAGAGCTG-----AACGGGCGAG 2448

QY      84  SerIleThrAsnValGlnGlyArgAspIleHisLysGlnHisLeuGlnGluLys 103
      |||||  ::::|  ::::|  |||||  |||||  |||||  |||||  |||||  |||||  ::
DB      2449  ATCAATTACCTCGAGATCCAGAGGCGCAAGAGCCTCTTCACAGCCTTCTTGAGTCC 2508

QY      104  LysAspGlySerIleLysProGlnGlnLysGluAsp 115
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ::
DB      2509  CTGGCTGAGAGATCAGCTCCGCTCCGAGATGAG 2544
```

Search completed: December 7, 2005, 16:14:58
Job time : 414.905 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 12:39:12 ; Search time 80.5219 Seconds
(without alignments)
543.218 Million cell updates/sec

Title: US-09-837-344-20
Perfect score: 603
Sequence: 1 LQEQQRDLQRKADTKNLR.....GHLEKKDGSIKPEQKEDKS 117

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3392430 seqs, 186927314 residues
Total number of hits satisfying chosen parameters: 6784660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.spool h/US09837344/runat.05122005.133535.865/app.query.fasta_1.718
-DB=Published Applications NA New -OPMT=fastap -SUFFIX=rnbn -MINMATCH=0.1
-LOOPC=0 -LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62
-TRANS=numan40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09837344 @CGN 1.1.231 @runat.05122005.133535.865
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSELOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:
1: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	86	14.3	1344	US-10-793-626-2899
2	86	14.3	1353	US-10-793-626-3225
3	86	14.3	3001	US-10-793-626-3813
4	86	14.3	3974	US-10-793-626-4245
5	82.5	13.7	193084	US-11-121-086-82
6	79.5	13.2	8424	US-10-821-234-47
7	79	13.1	423	US-10-793-626-2267
8	79	13.1	423	US-10-793-626-2437

9	79	13.1	2525	6	US-10-770-726-5	Sequence 5, Appl1
10	79	13.1	3019	6	US-10-793-626-4232	Sequence 4232, Ap
11	79	13.1	3726	6	US-10-793-626-3917	Sequence 3917, Ap
12	79	13.1	3986	6	US-10-793-626-4014	Sequence 4014, Ap
13	78.5	13.0	2040	6	US-10-750-185-41854	Sequence 41854, A
14	77.5	12.9	120697	7	US-11-121-086-48	Sequence 48, Appl
15	77	12.8	173602	7	US-11-121-086-25	Sequence 25, Appl
16	75	12.4	1458	6	US-10-508-263-21	Sequence 21, Appl
17	75	12.4	8948	6	US-10-623-155-119	Sequence 119, Ap
18	75	12.4	150468	7	US-11-112-908-56	Sequence 56, Appl
19	75	12.4	193789	7	US-11-112-908-55	Sequence 55, Appl
20	74	12.3	98862	7	US-11-121-086-76	Sequence 76, Appl
21	72.5	12.0	1476	6	US-10-793-626-2807	Sequence 2807, Ap
22	72.5	12.0	3857	6	US-10-793-626-480	Sequence 3480, Ap
23	71.5	11.9	636	6	US-10-467-657-5485	Sequence 5485, Ap
24	71.5	11.9	636	6	US-10-467-657-7403	Sequence 7403, Ap
25	71.5	11.9	648	6	US-10-793-626-595	Sequence 595, App
26	71.5	11.9	648	6	US-10-793-626-2257	Sequence 2257, Ap
27	71	11.8	984	6	US-10-467-657-6641	Sequence 6641, Ap
28	71	11.8	984	6	US-10-467-657-7709	Sequence 7709, Ap
29	71	11.8	1605	6	US-10-750-185-60116	Sequence 60116, A
30	71	11.8	98862	7	US-11-121-086-76	Sequence 76, Appl
31	70.5	11.7	451	7	US-11-108-172-1038	Sequence 1038, Ap
32	70.5	11.7	1200	6	US-10-131-826A-273	Sequence 273, App
33	70.5	11.7	2158	6	US-10-909-125-805	Sequence 805, App
34	70.5	11.7	149419	7	US-11-112-908-49	Sequence 49, Appl
35	70.5	11.7	161726	7	US-11-112-908-48	Sequence 48, Appl
36	70.5	11.7	161726	7	US-11-112-908-52	Sequence 52, Appl
37	70.5	11.7	166111	7	US-11-112-908-47	Sequence 47, Appl
38	70	11.6	762	6	US-10-467-657-3449	Sequence 3449, Ap
39	69	11.4	2845	6	US-10-909-125-1740	Sequence 1740, Ap
40	69	11.4	5233	7	US-11-060-914-1	Sequence 1, Appl
41	69	11.4	162537	7	US-11-121-086-59	Sequence 59, Appl
42	68.5	11.4	207835	7	US-11-121-086-59	Sequence 39, Appl
43	68.5	11.4	207835	7	US-11-121-086-40	Sequence 40, Appl
44	68	11.3	821	6	US-10-750-185-46312	Sequence 46312, A
45	68	11.3	2583	6	US-10-750-185-37927	Sequence 37927, A

ALIGNMENTS

RESULT 1
US-10-793-626-2899
; Sequence 2899, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P33480US
; CURRENT APPLICATION NUMBER: US/10/793, 626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164, 258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2899
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-2899

Alignment Scores:

Pred. No.:	1.21	Length:	1344
Score:	86.00	Matches:	29
Percent Similarity:	43.44%	Conservative:	24
Best Local Similarity:	23.77%	Mismatches:	43
Query Match:	14.26%	Indels:	26
DB:	6	Gaps:	4

US-09-837-344-20 (1-117) x US-10-793-626-2899 (1-1344)

```
QY      1 LeuGlnGlnGlnGlnArgAspLeuGlnGlnArgGlyAlaAspThrLysLysAsnLeuGlu 20
      43 ATTCTAAAGCGCAAGACAGTCAAAAAGCGCACTCGAAACCTGCTTAATTTAA 102
QY      21 -----ArgLysGlyHisGlyAspIleLeuAlaGlnAspLeuTyrgly 35
      103 GCTGATATTGATGCTCAAAAAGAAAGATGCAAGAACTC-----GAACAGCTTTCAAAA 156
QY      36 ArgLeuGlnIleProAlaIleGluLeuProSerGluAsnGluArgGlyTyrglyPro 55
      157 GAAATTGAAGCCTCAGCA-----CCT 177
QY      56 HisGlnSerSerLeuProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSer 75
      178 AAACAAGATGAACCACTTAAGATGAGTGCAGAACTTGAAGATTAACAAGATGTAAT 237
QY      76 IleIleGluLysThrAsnArgGlnSerIleThrThrAsnValGlnGlyArgAspIle 95
      238 TCTGAGAAAGAAATCAGAGAACAACTTGTATGATGAACAGAGAACTTCAGAT--- 294
QY      96 HisLysGlyHisLeuGlnGlnLysAspGlySerIleLysProGlnGlnLysGluAsp 115
      295 -----GAAGAAAACCTGATGATGACCAAAACAGATGACAAACCTGAA 339
QY      116 LysSer 117
      340 GAAACA 345
DB
RESULT 2
US-10-793-626-3225
/ Sequence 3225, Application US/10793626
/ Publication No. US20050255478A1
/ GENERAL INFORMATION:
/ APPLICANT: KIMBERLY, WILLIAM JOHN
/ TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
/ FILE REFERENCE: PU3480US
/ CURRENT APPLICATION NUMBER: US/10/793,626
/ PRIOR FILING DATE: 2004-03-04
/ PRIOR APPLICATION NUMBER: 60/164,258
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 4472
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 3225
/ LENGTH: 1353
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic
/ OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3225

Alignment Scores:
Pred. No.:      1 22      length:      1353
Score:          86.00     Matches:      29
Percent Similarity: 43.44%     Conservative: 24
Best Local Similarity: 23.77%     Mismatches: 43
Query Match:    14.26%     Indels:      26
DB:              6      Gaps:          4
US-09-837-344-20 (1-117) x US-10-793-626-3225 (1-1353)
QY      1 LeuGlnGlnGlnGlnArgAspLeuGlnGlnArgGlyAlaAspThrLysLysAsnLeuGlu 20
      43 ATTCTAAAGCGCAAGACAGTCAAAAAGCGCACTCGAAACCTGCTTAATTTAA 102
DB
QY      21 -----ArgLysGlyHisGlyAspIleLeuAlaGlnAspLeuTyrgly 35
      103 GCTGATATTGATGCTCAAAAAGAAAGATGCAAGAACTC-----GAACAGCTTTCAAAA 156
DB
QY      36 ArgLeuGlnIleProAlaIleGluLeuProSerGluAsnGluArgGlyTyrglyPro 55
      157 GAAATTGAAGCCTCAGCA-----CCT 177
DB
```

```
QY      56 HisGlnSerSerLeuProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSer 75
      178 AAACAAGATGAACCACTTAAGATGAGTGCAGAACTTGAAGATTAACAAGATGTAAT 237
DB
QY      76 IleIleGluLysThrAsnArgGlnSerIleThrThrAsnValGlnGlyArgAspIle 95
      238 TCTGAGAAAGAAATCAGAGAACAACTTGTATGATGAACAGAGAACTTCAGAT--- 294
QY      96 HisLysGlyHisLeuGlnGlnLysAspGlySerIleLysProGlnGlnLysGluAsp 115
      295 -----GAAGAAAACCTGATGATGACCAAAACAGATGACAAACCTGAA 339
QY      116 LysSer 117
      340 GAAACA 345
DB
RESULT 3
US-10-793-626-3813/c
/ Sequence 3813, Application US/10793626
/ Publication No. US20050255478A1
/ GENERAL INFORMATION:
/ APPLICANT: KIMBERLY, WILLIAM JOHN
/ TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
/ FILE REFERENCE: PU3480US
/ CURRENT APPLICATION NUMBER: US/10/793,626
/ PRIOR FILING DATE: 2004-03-04
/ PRIOR APPLICATION NUMBER: 60/164,258
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 4472
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 3813
/ LENGTH: 3001
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic
/ OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3813

Alignment Scores:
Pred. No.:      3 44      length:      3001
Score:          86.00     Matches:      29
Percent Similarity: 43.44%     Conservative: 24
Best Local Similarity: 23.77%     Mismatches: 43
Query Match:    14.26%     Indels:      26
DB:              6      Gaps:          4
US-09-837-344-20 (1-117) x US-10-793-626-3813 (1-3001)
QY      1 LeuGlnGlnGlnGlnArgAspLeuGlnGlnArgGlyAlaAspThrLysLysAsnLeuGlu 20
      2456 ATTCTAAAGCGCAAGACAGTCAAAAAGCGCACTCGAAACCTGCTTAATTTAA 2397
DB
QY      21 -----ArgLysGlyHisGlyAspIleLeuAlaGlnAspLeuTyrgly 35
      2396 GCTGATATTGATGCTCAAAAAGAAAGATGCAAGAACTC-----GAACAGCTTTCAAAA 2343
DB
QY      36 ArgLeuGlnIleProAlaIleGluLeuProSerGluAsnGluArgGlyTyrglyPro 55
      2342 GAAATTGAAGCCTCAGCA-----CCT 2322
DB
QY      56 HisGlnSerSerLeuProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSer 75
      2321 AAACAAGATGAACCACTTAAGATGAGTGCAGAACTTGAAGATTAACAAGATGTAAT 2262
DB
QY      76 IleIleGluLysThrAsnArgGlnSerIleThrThrAsnValGlnGlyArgAspIle 95
      2261 TCTGAGAAAGAAATCAGAGAACAACTTGTATGATGAACAGAGAACTTCAGAT--- 2205
DB
QY      96 HisLysGlyHisLeuGlnGlnLysAspGlySerIleLysProGlnGlnLysGluAsp 115
      2204 -----GAAGAAAACCTGATGATGACCAAAACAGATGACAAACCTGAA 2160
DB
```


QY 116 LysSer 117
Db 2159 GAAACA 2154

RESULT 4

US-10-793-626-4245
Sequence 4245, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P134805
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4245
LENGTH: 3974
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4245

Alignment Scores:
Pred. No.: 4.97 Length: 3974
Score: 86.00 Matches: 29
Percent Similarity: 43.44% Conservative: 24
Best Local Similarity: 23.77% Mismatches: 43
Query Match: 14.26% Indels: 26
DB: Gaps: 4

US-09-837-344-20 (1-117) x US-10-793-626-4245 (1-3974)

QY 1 LeuGlnGlnGlnGlnAArgApleuGlnGlnAArgLysAlaAspThrLysLysAsnLeuGlu 20
Db 460 ATTCTAAGCGCAAGAAAGACAGTCGAAAGGCGACCTCGAAACCTGCTGTAATTAAAA 519
QY 21 -----AArgLysLysGlnGlnHsGlnLysAspLysLeuAlaGlnAspLeuTyGly 35
Db 520 GCTGATATGTGATGCTCAAAAGAAAGAGTACGAAAGAACCTC-----GAACAGCTTCAAAA 573
QY 36 ArgLeuGlnLysProAlaLysLeuProSerGlnAsnGlnArgLysTyTyrLysPro 55
Db 574 GAATTGAAAGCGTCAGCA-----CCT 594
QY 56 HisGlnSerSerLeuProGlnAspAsnArgLysAsnSerArgAspSerLysGlnLysSer 75
Db 595 AAACAGATGAAACCACTAAAGATGAAGCTCAGACAGTGAAGTGAACAAAGATGATAT 654
QY 76 LLeuGlnLysThrAsnArgLysSerLysThrAsnValGlnGlnArgArgAspLys 95
Db 655 TCTGAGAGAAATGACGAAACAACTTCTGATGATGAACCAAGAGAACTTCAGAT--- 711
QY 96 HisLysGlnLysLeuGlnGlnLysLysAspGlnLysLysProGlnGlnLysGlnAsp 115
Db 712 -----GAAAGAAACCTGATGATGACCAAAACAGATGACAAACCTGAA 756
QY 116 LysSer 117
Db 757 GAAACA 762

RESULT 5

US-11-121-086-82
Sequence 82, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KRISTEN V.

TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138.6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 82
LENGTH: 193084
TYPE: DNA
ORGANISM: Homo sapiens

US-11-121-086-82

Alignment Scores:
Pred. No.: 2.12e+03 Length: 193084
Score: 82.50 Matches: 35
Percent Similarity: 43.18% Conservative: 22
Best Local Similarity: 26.52% Mismatches: 44
Query Match: 13.68% Indels: 31
DB: Gaps: 7

US-09-837-344-20 (1-117) x US-11-121-086-82 (1-193084)

QY 1 LeuGlnGlnGlnGlnAArgApleuGlnGlnAArgLysAlaAspThrLysLysAsnLeuGlu 20
Db 137155 ATATCAGAGATTGAAAGTGAATCTTAATGAATAGAGAAAGCAAAATTAAGAAAA 137214
QY 21 ArgLysLysGlnGlnHsGlnLysAspLysLeuAlaGln-----AspLeuTyGlnArgLeuGlu 38
Db 137215 AGAATGAAAGAAACAACAAGAGCTCCAAAGAAATATGGACTATGCAAAAGCAAAAT 137274
QY 39 LLeuProAlaLysLeuLysProSer-----GlnAsnGlu 49
Db 137275 TTACGTTTGAATGGAATATCACTGAAGTGAAGAGAAAGCAAGCAAGTGGAAACTCT 137334
QY 50 ArgGlnTyTyrLysProHisGlnSerSerLeuProGln----- 62
Db 137335 TCAGGATATAT-----CCAGGAAACTTCACAACTTGAAGAGACAGGCCAACAT 137385
QY 63 AspAsnArgLysAsnSerArgAspSerLysGlnLysSerLysLeuGlnLysThrAsnArg 82
Db 137386 TCAATTAAGGAATATACGAAAGAACCAACAAGTA---CTCCGAGAGAGAGCAACCC 137442
QY 83 GluSerLysThrAsnValGlnGlnArgArgAspLysHisLysGlnLysLeuGlnGlu 102
Db 137443 AAGACA-----CATATCATGATGATTCACCAAGGTCAAAATGAAGGA 137484
QY 103 LysLysAspGlnSerLysLysProGlnGlnLysGln 114
Db 137485 AAA-----AATGTTAAGGCGACCCAGAGAGAA 137511

US-10-821-234-47

Sequence 47, Application US/10821234
Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarmani, Susan
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821,234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: PL_SEQ_genes Version 1.0
SEQ ID NO 47
LENGTH: 8424
TYPE: DNA
ORGANISM: Homo sapiens

US-10-821-234-47

Alignment Scores:

Alignment No.:	82.8	Length:	8424
Score:	79.50	Matches:	29
Percent Similarity:	42.20%	Conservative:	17
Best Local Similarity:	26.61%	Mismatches:	38
Query Match:	13.18%	Indels:	25
DB:	6	Gaps:	5

US-09-837-344-20 (1-117) x US-10-821-234-47 (1-8424)

QY 17 LysAsnLeuGluArgLysLeu-----HisGlyAspIleLeuIleGlu 31

DB 5267 AAGGACATGAAAGCTTTCAAGACCTTAGAGCTGCATTGAAGAGACCTTACTTACCTGCAAGT 5326

QY 32 AspLeuTyrGlyArgLeuGluIleProAlaIleGluLeuProSerGluAsnGluArgGly 51

DB 5327 GATGTT--GGAAGT-----GACAGATCT 5347

QY 52 TyrTyrIleProHisGlnSerLeuProGln-----AspAsnArgGlyAsnSer 68

DB 5348 GCTGCCAGCCCGGTTGTAAGTATGCGAAGAAAGAGCTTCAGAGCTTTCAGAGGAA 5407

QY 69 ArgAspSerLysGluIleSerIleIleGluLysThrAsnArgGluSerIleThrAsn 88

DB 5408 AAGATGATTATGAAATTTTGTAAAGTTAAGGACACTCAGCAAAAGCAAGAAAT 5467

QY 89 ValGluGlyArgArgAspIleHisLysGlyHisLeuGluGluLysAspGlySerIle 108

DB 5468 -----AAGAACCCTGATTAAGGGGAGAAAGAGAAAGAGACTTCTATT 5515

QY 109 LysProGluGlnLysGluAspLysSer 117

DB 5516 AGATCTCGAAGTAAAGCTTCCAATCT 5542

RESULT 7

US-10-793-626-2267

Sequence 2267, Application US/10793626

Publication No. US20050255478A1

GENERAL INFORMATION:

APPLICANT: KIMBERLY, WILLIAM JOHN

TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

FILE REFERENCE: PU3480US

CURRENT APPLICATION NUMBER: US/10/793,626

CURRENT FILING DATE: 2004-03-04

PRIOR APPLICATION NUMBER: 60/164,258

PRIOR FILING DATE: 1999-11-09

NUMBER OF SEQ ID NOS: 4472

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2267

LENGTH: 423

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

OTHER INFORMATION: nucleic acid sequence

US-10-793-626-2267

Alignment Scores:

Pred. No.:	1.92	Length:	423
Score:	79.00	Matches:	28
Percent Similarity:	46.40%	Conservative:	30
Best Local Similarity:	22.40%	Mismatches:	35
Query Match:	13.10%	Indels:	32
DB:	6	Gaps:	6

US-09-837-344-20 (1-117) x US-10-793-626-2267 (1-423)

QY 1 LeuGlnGluGlnGlnArgAspLeuGluGlnArgLysAlaAspThrLysLysAsnLeuGlu 20

DB 100 TTGAAAAAGAAATTAAGATTTTAAAAAACAAGAAAGAAAGTTACAGAGCAAAAGAT 159

QY 21 ArgLysLysGluHisGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGluIlePro 40

DB 160 AAGCTTAACACAAACAGATAGCTTCCAGAAAGATGTAAT----- 201

QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHisGlnSerLeu 60

DB 202 -----GACTTGCCCTTAAGAAC-----ACATCCCGA 228

QY 61 ProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThr 80

DB 229 GATTAAGAAAAATTAAGATTAAT--CATATGCAAAAGAAAGCTTCAGATATATCATGC 285

QY 81 -----AsnArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHis--- 96

DB 286 ACATGCTGTAATCATGATGATCAAACTTAACAAATTAAGCAATCAAGATGACATGAC 345

QY 97 -----LysGlyHisLeuGluGluLysLysAspGlySerIleLysProGluGln 112

DB 346 AGTCAATCCTTAACCAATACATACAGCAG-----AAGCCTTCACAG 387

QY 113 LysGluAspLysSer 117

DB 388 AATGATAGAAAAAT 402

RESULT 8

US-10-793-626-2437

Sequence 2437, Application US/10793626

Publication No. US20050255478A1

GENERAL INFORMATION:

APPLICANT: KIMBERLY, WILLIAM JOHN

TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

FILE REFERENCE: PU480US

CURRENT APPLICATION NUMBER: US/10/793,626

CURRENT FILING DATE: 2004-03-04

PRIOR APPLICATION NUMBER: 60/164,258

PRIOR FILING DATE: 1999-11-09

NUMBER OF SEQ ID NOS: 4472

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2437

LENGTH: 423

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

OTHER INFORMATION: nucleic acid sequence

US-10-793-626-2437

Alignment Scores:

Pred. No.:	1.92	Length:	423
Score:	79.00	Matches:	28
Percent Similarity:	46.40%	Conservative:	30
Best Local Similarity:	22.40%	Mismatches:	35
Query Match:	13.10%	Indels:	32
DB:	6	Gaps:	6

US-09-837-344-20 (1-117) x US-10-793-626-2437 (1-423)

QY 1 LeuGlnGluGlnGlnArgAspLeuGluGlnArgLysAlaAspThrLysLysAsnLeuGlu 20

DB 100 TTGAAAAAGAAATTAAGATTTTAAAAAACAAGAAAGAAAGTTACAGAGCAAAAGAT 159

QY 21 ArgLysLysGluHisGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGluIlePro 40

DB 160 AAGCTTAACACAAACAGATAGCTTCCAGAAAGATGTAAT----- 201

QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHisGlnSerLeu 60

DB 202 -----GACTTGCCCTTAAGAAC-----ACATCCCGA 228

QY 61 ProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThr 80

DB 229 GATTAAGAAAAATTAAGATTAAT--CATGATGCAAAAGAAAGCTTCAGATATATCATGC 285

```

Oy      81 -----AsnArgLuserlethrhrAsenValgluglyArGArGAepLIeH's--- 96
        ||| :||| |||
Db      286 ACATCTGCTAAATCATGATGATCAAACTAACAAAATAAAGACATCAAGATGAACAATGCAC 345
        ||| :||| |||
Oy      97 -----LygLYHIshLeuGlugluYslYsaerGLyserIleLyseProGIugIn 112
        ||| :||| |||
Db      346 AGTCAATCCTCTAAACCAACTACACAGCAG-----AACCCCTCACAG 387
        ||| :||| |||
Oy      113 LyegLiuaepLyser 117
        ::: |||:::
Db      388 AATGATAGAAAAAAT 402

RESULT 9
US-10-770-726-5
; Sequence 5, Application US/10770726
; Publication No. US20050266409a1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: CANCERS
; FILE REFERENCE: AM101079 (031896-010000)
; CURRENT APPLICATION NUMBER: US/10/770,726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 2525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-770-726-5
```

[illegible]

```

Sequence 4232, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4232
LENGTH: 3019
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4232

```

[illegible]

Length:	35.9	3986
Matches:	79.00	28

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 04:22:54 ; Search time 1139.77 Seconds
(without alignments)
8665.841 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 1482

Sequence: 1 CAAGAAACAACAAGCATCT.....ATATTATGAACTATA 1482

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_21:*

- 1: geneeqn1980s:*
- 2: geneeqn1990s:*
- 3: geneeqn2000s:*
- 4: geneeqn2001s:*
- 5: geneeqn2002s:*
- 6: geneeqn2003s:*
- 7: geneeqn2004s:*
- 8: geneeqn2005s:*
- 9: geneeqn2006s:*
- 10: geneeqn2007s:*
- 11: geneeqn2008s:*
- 12: geneeqn2009s:*
- 13: geneeqn2010s:*
- 14: geneeqn2011s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1482	100.0	1496	2	AAQ28119
2	950	64.1	950	2	AAQ28115
3	853.2	57.6	1320	2	AAQ80916
4	586.4	39.6	1566	2	AAQ05140
5	503.4	34.0	1371	12	ADO21941
6	493.6	33.3	1374	12	ADO21919
7	472.8	31.9	988	2	AAQ28117
8	233.4	15.7	1300	12	ADP85917
9	225.8	15.2	3399	2	AAQ05868
10	183.2	12.4	210	1	AAH80899
11	180.2	12.2	1000	12	ADQ62833
12	180.2	12.2	1000	12	ADQ62832
13	179.4	12.1	1686	2	AAQ87587
14	176.6	11.9	3579	3	AAH70099
15	170.6	11.5	1998	3	AAH70212
16	151.2	10.2	6292	4	AAH46735
17	149.2	10.1	6767	4	AAH46608
18	146	9.9	3662	4	ABJ05274
19	139.4	9.4	9539	4	AAH45347

C	20	139.4	9.4	9539	6	ABK28180
C	21	137.8	9.3	6668	6	ABJ33697
C	22	136.4	9.2	6644	2	AAH33181
C	23	136.4	9.2	7372	2	AAH33182
C	24	136.4	9.2	7797	2	AAH33180
C	25	136.4	9.2	7996	2	AAH33184
C	26	134.2	9.1	5493	13	ADH89675
C	27	133.8	9.0	778	6	ABJ15588
C	28	132.8	9.0	778	6	ABJ15589
C	29	131.4	8.9	5493	13	ADH89401
C	30	131.4	8.9	7442	4	AAH46686
C	31	129.8	8.8	1039	14	ACI64817
C	32	129.8	8.8	1762	6	ABQ20939
C	33	129.8	8.8	1762	6	ABQ20938
C	34	129.2	8.7	1995	6	ABQ20942
C	35	129.2	8.7	1995	6	ABQ20943
C	36	129	8.7	1200	6	ABQ39210
C	37	129	8.7	1200	6	ABQ39211
C	38	129	8.7	7597	6	ABJ33013
C	39	129	8.7	14006	6	ABJ33958
C	40	127.4	8.6	3683	8	ABJ10199
C	41	127	8.6	3683	8	ABJ10053
C	42	123.4	8.3	3037	6	ABQ54066
C	43	123.4	8.3	3037	6	ABQ54067
C	44	123	8.3	12337	6	ABJ34358
C	45	120.8	8.2	612	14	ACI64894

ALIGNMENTS

RESULT 1	AAQ28119	standard; DNA; 1496 BP.
ID	AAQ28119	
XX	AAQ28119;	
AC	25-MAR-2003	(revised)
DT	08-FEB-1993	(first entry)
XX		
DE	P. falciparum	LSA gene 3' region.
XX		
KM	Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope; paludism; liver stage-specific antigen; ss.	
OS	Plasmodium falciparum.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	1..1482
FT		/*tag= C
FT	repeat_region	/product= "LSA_C-terminal_region"
FT		/tag= a
FT	repeat_unit	37..87
FT		/*tag= b
FT		/type= TANDEM
XX		
XX	WO9213884-A1.	
XX	20-AUG-1992.	
XX	05-FEB-1992;	92WO-FR000104.
XX	05-FEB-1991;	91FR-00001286.
XX		
XX	(INSP) INST PASTEUR.	
XX	Guerinmarchand C, Druilhe P;	
XX	WPI; 1992-299985/36.	
XX	P-PSDB; AAR26944.	
XX		
XX	Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for	

ABK28180 DNA trans
ABJ33697 Human imm
AAH33181 Base sequ
AAH33182 Base sequ
AAH33180 Compox vi
AAH33184 Base sequ
ADH89675 Oligonuc
ABJ15588 Oligonuc
ABJ15589 Oligonuc
ADH89401 Oligonuc
AAH46686 Tumour su
ACI64817 M. xanthu
ABQ20939 Oligonuc
ABQ20942 Oligonuc
ABQ20943 Oligonuc
ABQ39210 Oligonuc
ABQ39211 Oligonuc
ABJ33013 Human imm
ABJ33958 Human imm
ABJ10199 Haematopo
ABJ10053 Haematopo
ABQ54066 Oligonuc
ABQ54067 Oligonuc
ABJ34358 Human imm
ACI64894 M. xanthu

PT vaccination against, treatment of and diagnosis of malaria.
XX
PS Claim 2; Fig 8-10; 81pp; French.
XX

CC This nucleotide sequence is the 3' part of the P.falciparum liver-stage
CC specific antigen (LSA) gene. It codes for a polypeptide sequence which
CC carries a T cell epitope characteristic of a protein produced in
CC hepatocytes infected with P.falciparum. The polypeptide can be used in
CC the preparation of vaccines against malaria. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX

Sequence 1496 BP; 716 A; 169 C; 300 G; 311 T; 0 U; 0 Other;

Query Match 100.0%; Score 1482; DB 2; Length 1496;

Best Local Similarity 100.0%; Pred. No. 8.5e-202;

Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CAAAGAACAAACAGCATCTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAGAACAA 60
DB 1 CAAAGAACAAACAGCATCTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAGAACAA 60
QY 61 CAAAGCATTTAGAACAGATGACTTGTCTAAAGAAAAGTTGCAGAGCAAGCGAT 120
DB 61 CAAAGCATTTAGAACAGATGACTTGTCTAAAGAAAAGTTGCAGAGCAAGCGAT 120
QY 121 TTAGAACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAACAAACAGCATCTAGAACAA 180
DB 121 TTAGAACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAACAAACAGCATCTAGAACAA 180
QY 181 GAGAGACGCTGTAAAGAAAAGTTGCAGAACAAACAGCATTTAGAACAGAGAGAGCT 240
DB 181 GAGAGACGCTGTAAAGAAAAGTTGCAGAACAAACAGCATTTAGAACAGAGAGAGCT 240
QY 241 GCTAAAGAAAAGTTGCAGAACAAACAGCATTTAGAACAGATGACTTGTCTAAAGAA 300
DB 241 GCTAAAGAAAAGTTGCAGAACAAACAGCATTTAGAACAGATGACTTGTCTAAAGAA 300
QY 301 AAGTTACAAGAGCAGCAAAAGCATTTAGAACAGAGAGCGTCTAAAGAAAAGTTGCA 360
DB 301 AAGTTACAAGAGCAGCAAAAGCATTTAGAACAGAGAGCGTCTAAAGAAAAGTTGCA 360
QY 361 GAAACAACAAAGCGATTTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAACAA 420
DB 361 GAAACAACAAAGCGATTTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAACAA 420
QY 421 AGCGATTTAGAACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAACAAACAGCATTTA 480
DB 421 AGCGATTTAGAACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAACAAACAGCATTTA 480
QY 481 GAAACAAGAGAGCGTCTAAAGAAAAGTTGCAGAACAAACAGCATTTAGAACAGAG 540
DB 481 GAAACAAGAGAGCGTCTAAAGAAAAGTTGCAGAACAAACAGCATTTAGAACAGAG 540
QY 541 AGACGCTGTAAAGAAAAGTTGCAGAACAAACAGCATTTAGAACAGAGAGAGCTGCT 600
DB 541 AGACGCTGTAAAGAAAAGTTGCAGAACAAACAGCATTTAGAACAGAGAGAGCTGCT 600
QY 601 AAAAGAAAAGTTGCAGAGCAGCAAGAGATTTAGAACAAAGAGAGCTGATCGAAAAA 660
DB 601 AAAAGAAAAGTTGCAGAGCAGCAAGAGATTTAGAACAAAGAGAGCTGATCGAAAAA 660
QY 661 AATTTCAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 720
DB 661 AATTTCAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 720
QY 721 GAAATACAGAGATGAGAACTTCATCAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 780
DB 721 GAAATACAGAGATGAGAACTTCATCAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 780
QY 781 TCTTCTTTAAGCTCAGAGCAACAGAGAGAAATAGTAGAGATTCCAGAGAAATATCTAATA 840
DB 781 TCTTCTTTAAGCTCAGAGCAACAGAGAGAAATAGTAGAGATTCCAGAGAAATATCTAATA 840
```

```
QY 841 GAAAAAACAAATAGAGATCTATTACAAACAATGTTGAGAGCAGAGGATATACATAAA 900
DB 841 GAAAAAACAAATAGAGATCTATTACAAACAATGTTGAGAGCAGAGGATATACATAAA 900
QY 901 GGACATCTTGAGAGAAAAGAAAAGATGTTCAATTAACCCAGAACAAAAGAAATAACT 960
DB 901 GGACATCTTGAGAGAAAAGAAAAGATGTTCAATTAACCCAGAACAAAAGAAATAACT 960
QY 961 GCTGACATCAAAATCATCTACCTTAGACACAGTAATATTTCTGATGTTAAGATTTTCAA 1020
DB 961 GCTGACATCAAAATCATCTACCTTAGACACAGTAATATTTCTGATGTTAAGATTTTCAA 1020
QY 1021 ATTAAGTAAATAGATGAGATGAAATAGTGCCTAATATGATGATGATGATGATGATG 1080
DB 1021 ATTAAGTAAATAGATGAGATGAAATAGTGCCTAATATGATGATGATGATGATGATG 1080
QY 1081 GAAAGATGAGAGACTTAAAGCAATTTAAGCTTATGCTAATATGCAATTTCCAGAT 1140
DB 1081 GAAAGATGAGAGACTTAAAGCAATTTAAGCTTATGCTAATATGCAATTTCCAGAT 1140
QY 1141 GAAAGAAACATAGCAATTTATTAAGAACTAGAAATTTGATAGAGAAAAGAAAATTTA 1200
DB 1141 GAAAGAAACATAGCAATTTATTAAGAACTAGAAATTTGATAGAGAAAAGAAAATTTA 1200
QY 1201 GATGATTTAGATGAGAGAAATGAGAAAATCATCAGAGAAATATGCAAGAAAATATAA 1260
DB 1201 GATGATTTAGATGAGAGAAATGAGAAAATCATCAGAGAAATATGCAAGAAAATATAA 1260
QY 1261 AAAGAAAAGAAATATGAGAAAAGAAAAGAAAGATTAATTTTAAACCAATGATTAAGTTG 1320
DB 1261 AAAGAAAAGAAATATGAGAAAAGAAAAGAAAGATTAATTTTAAACCAATGATTAAGTTG 1320
QY 1321 TATGATGAGCATTTATTAAGAAATATTAAGCAAGTTATTAAGAGAAAAGAAAAG 1380
DB 1321 TATGATGAGCATTTATTAAGAAATATTAAGCAAGTTATTAAGAGAAAAGAAAAG 1380
QY 1381 TTCAATTAATATCATTTGTTTATATTTGACGAGCAATGAAATTTTACAGATGCTGAT 1440
DB 1381 TTCAATTAATATCATTTGTTTATATTTGACGAGCAATGAAATTTTACAGATGCTGAT 1440
QY 1441 GAGTTATCTGAAGATATTAATTAATTTTATGAACATATTA 1482
DB 1441 GAGTTATCTGAAGATATTAATTAATTTTATGAACATATTA 1482
RESULT 2
AAQ28115
ID AAQ28115 standard; DNA; 950 BP.
XX
AC AAQ28115;
XX
DT 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
DE P. falciparum LSA-R-NR coding sequence.
XX
KW Malaria; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;
KW paludism; liver stage-specific antigen; ss.
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT repeat_region 3..629
FT FT /*tag= a
FT FT /rpt_type= TANDEM
FT repeat_unit 3..53
FT FT /*tag= b
XX
PN WO9213884-A1.
XX
PD 20-AUG-1992.
XX
PF 05-FEB-1992; 92WO-FR000104.
```



```

XX PR 05-FEB-1991; 91FR-00001286.
XX
XX (INSP) INSTR PASTEUR.
XX
XX Guertinmarchand C, Druilhe P;
XX
XX WPI; 1992-299985/36.
XX DR P-PSDB; AAR26941.
XX
XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX PT vaccination against, treatment of and diagnosis of malaria.
XX
XX Claim 17, Fig 2, 81pp; French.
XX
XX A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
XX CC gIII was used to transform E.coli. The expression library was screened
XX CC with human antisera against antigens of all stages of P. falciparum. The
XX CC library was reseeded with antibodies affinity-purified on a clone
XX CC which was able to recognise antibodies specific to the hepatic phase.
XX CC About 40 clones were detected which produced a characteristic LSA
XX CC epitope. The clone with the largest insert (950 bases) encoded LSA-R-NR
XX CC containing a 12-repeat region followed by a non-repeat region. Preferred
XX CC antigenic polypeptides from the invention are derived from the amino acid
XX CC sequence of LSA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 950 BP; 456 A; 127 C; 206 G; 161 T; 0 U; 0 Other;
XX
XX Query Match 64.1%; Score 950; DB 2; Length 950;
XX Best Local Similarity 100.0%; Pred. No. 2.9e-126;
XX Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 11 AAGCGATTTCAGACAAAGAGACGCTCTAAGAAAAGTTGCAAGAACAAAGCGATT 70
Db 1 AAGCGATTTCAGACAAAGAGACGCTCTAAGAAAAGTTGCAAGAACAAAGCGATT 60
71 TAGAACAAAGATAGACTTGTCTAAGAAAAGTTACAGAGCAGCAAGCGATTTCAGACAAAG 130
61 TAGAACAAAGATAGACTTGTCTAAGAAAAGTTACAGAGCAGCAAGCGATTTCAGACAAAG 120
131 AGAGACTTGTCTAAGAAAAGTTGCAAGAACAAAGCGATTTCAGACAAAGAGACGCTG 190
Db 121 AGAGACTTGTCTAAGAAAAGTTGCAAGAACAAAGCGATTTCAGACAAAGAGACGCTG 180
QY 191 CTAAAGAAAAGTTGCAAGAACAAAGCGATTTCAGACAAAGAGACGCTGCTAAAGAAA 250
Db 181 CTAAAGAAAAGTTGCAAGAACAAAGCGATTTCAGACAAAGAGACGCTGCTAAAGAAA 240
251 AGTTGCAAGAACAAAGCGATTTCAGACAAAGATAGACTTGTCTAAGAAAAGTTACAG 310
Db 241 AGTTGCAAGAACAAAGCGATTTCAGACAAAGATAGACTTGTCTAAGAAAAGTTACAG 300
311 AGCAGCAAGCGATTTCAGACAAAGAGACGCTGCTAAGAAAAGTTGCAAGAACAAAG 370
Db 301 AGCAGCAAGCGATTTCAGACAAAGAGACGCTGCTAAGAAAAGTTGCAAGAACAAAG 360
QY 371 GCGATTTCAGACAAAGAGACGCTGCTAAGAAAAGTTGCAAGAACAAAGCGATTTCAG 430
Db 361 GCGATTTCAGACAAAGAGACGCTGCTAAGAAAAGTTGCAAGAACAAAGCGATTTCAG 420
431 AACCAAGAGAGACTTGTCTAAGAAAAGTTGCAAGAACAAAGCGATTTCAGACAAAG 490
Db 421 AACCAAGAGAGACTTGTCTAAGAAAAGTTGCAAGAACAAAGCGATTTCAGACAAAG 480
491 GACGTGCTAAGAAAAGTTGCAAGAACAAAGCGATTTCAGACAAAGAGACGCTGCTA 550
Db 481 GACGTGCTAAGAAAAGTTGCAAGAACAAAGCGATTTCAGACAAAGAGACGCTGCTA 540
QY 551 AAGAAAAGTTGCAAGAACAAAGCGATTTCAGACAAAGAGACGCTGCTAAGAAAAGT 610
Db 541 AAGAAAAGTTGCAAGAACAAAGCGATTTCAGACAAAGAGACGCTGCTAAGAAAAGT 600
QY 611 TGCAGAGAGCAGCAAGAGATTTCAGACAAAGAGCGCTGATTCGAAAAAAATTTAGAAA 670

```

```

Db 601 TGCAGAGAGCAGCAAGAGATTTCAGACAAAGAGCGCTGATTCGAAAAAAATTTAGAAA 660
QY 671 GAAAAAAGCAATGAGATATATTTCAGACAGAGATTTCATATGCTGTTAGAAATACAG 730
Db 661 GAAAAAAGCAATGAGATATATTTCAGACAGAGATTTCATATGCTGTTAGAAATACAG 720
QY 731 CTATGAACTTCATTCAGAAAATGAAACCTGATATATTATACCATCAATCTTCTTAC 790
Db 721 CTATGAACTTCATTCAGAAAATGAAACCTGATATATTATACCATCAATCTTCTTAC 780
QY 791 CTCAGAGCAACAGAGGAAATAGATTCAGATTCCAGAAATATCTTAATAGAAAAACAA 850
Db 781 CTCAGAGCAACAGAGGAAATAGATTCAGATTCCAGAAATATCTTAATAGAAAAACAA 840
QY 851 ATAGGAATCTATTTCACAAATGTTGAAGACGAAGGATATACATAAGACATCTTG 910
Db 841 ATAGGAATCTATTTCACAAATGTTGAAGACGAAGGATATACATAAGACATCTTG 900
QY 911 AAGAAAAGAAATGCTTCATATTAACCGAACCAAAAAAGAAAGATTAATCT 960
Db 901 AAGAAAAGAAATGCTTCATATTAACCGAACCAAAAAAGAAAGATTAATCT 950
RESULT 3
AA080916
ID AA080916 standard; cDNA; 1320 BP.
XX
XX AA080916;
XX AC
XX 16-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 29-AUG-1995 (first entry)
XX DE
XX DE Plasmodium falciparum LSA-1 repeatless gene cDNA.
XX
XX KW Plasmodium falciparum LSA-1 repeatless gene; recombinant poxvirus;
XX KW multicomponent multistage malarial vaccines; immunogens;
XX KW malaria diagnosis; ss.
XX
XX OS Plasmodium falciparum; (PLSARPLS.14L.1).
XX
XX PN WO9428930-A1.
XX
XX PD 22-DEC-1994.
XX PF 10-JUN-1994; 94WO-US006652.
XX
XX PR 11-JUN-1993; 93US-00075783.
XX PR 09-JUN-1994; 94US-00257073.
XX
XX PA (VIRO-) VIROGENETICS CORP.
XX
XX PI Paolletti E, De Taisne C, Tine JA;
XX PT WPI; 1995-036113/05.
XX PT Recombinant poxvirus contg. Plasmodium DNA in non-essential region -
XX PT useful in vaccines against malaria and for prodn. of Plasmodium
XX PT immunogens.
XX
XX PS Claim 3; Fig 11; 183pp; English.
XX
XX AA080916 is the P. falciparum LSA-1 repeatless gene cDNA sequence. New
XX CC recombinant poxviruses containing either the SPRA, ABRA, pfhap70, AMA-1,
XX CC Pf825, Pf816, CSP, PfSS92, LSA-1, LSA-1 repeatless, MSA-1, MSA-1 (N-
XX CC terminal p83 or C-terminal gp42) genes, or a combination of these in non-
XX CC essential regions of their genomes are claimed. These poxviruses (pref.
XX CC with a virulence reducing genomic deletion or disruption) can be used as
XX CC vaccines against malaria and for the prodn. of Plasmodium immunogens.
XX CC These viruses provide multicomponent, multistage vaccines due to their
XX CC expression of sporozoite, liver stage, blood stage and sexual stage
XX CC proteins. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-

```

CC OCT-2003 to standardise OS field)

XX Sequence 1320 BP; 619 A; 128 C; 227 G; 346 T; 0 U; 0 Other;

Query Match 57.6%; Score 853.2; DB 2; Length 1320;
 Best Local Similarity 87.9%; Pred. No. 1.5e-112;
 Matches 930; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

```

QY 425 ATTGAAACAAGAGAGCTTGTAAAGAAAGTTGCAAGACAAACAGATTAGAAC 484
DB 263 ATGTAAAAATGTGTACAAACAAATTTCAAAAGCTTTTAAAGAACTTGTGTTTCKG 322
QY 485 AAGAGAGCGTCTAAAGAAAGTTGCAAGAAACAAACAAACGATTAGAACAGAGAGC 544
DB 323 AGAATATATTCCTTAAAGAAATTAATTAATAGAGAGGAAATTAAATGAAACATTA 382
QY 545 GTGCTAAAGAAAGTTGCAAGAAACAAACGATTAGAACAGAGAGAGCTGCTAAG 604
DB 383 TAAATGATGATGACGATTAAGAAATATATTAAAGGCGACAGAAACAGACAAAG 442
QY 605 AAAAGTTGCAAGAGCAGCAAGAGATTGTAAACAAAGAGGCTGATAGAAAAAATT 664
DB 443 ATCTTGAAAGAAAGAGCGCGCCTTAAGAAACAGAGAGGCTGATAGAAAAAATT 502
QY 665 TAGAAAGAAAGAAAGCAATGAGATATATTAGCAGAGATTATATGCTTTAGAA 724
DB 503 TAGAAAGAAAGAAAGCAATGAGATGATTTAGCAGAGATTATATGCTTTAGAA 562
QY 725 TACAGCTATAGACTTCCATCAGAAATGAAAGTGGATATTATATACCATCATCTT 784
DB 563 TACAGCTATAGACTTCCATCAGAAATGAAAGTGGATATTATATACCATCATCTT 622
QY 785 CTTTACCTCAGAGCAACAGAGGAAATAGAGATTCCAAGAAATATCTATATAGAA 844
DB 623 CTTTACCTCAGAGCAACAGAGGAAATAGAGATTCCAAGAAATATCTATATAGAA 682
QY 845 AAACAATAGAGATCTATTACACAAATGTTGAAGACGAAGGATATATACATAAGAC 904
DB 683 AAACAATAGAGATCTATTACACAAATGTTGAAGACGAAGGATATATACATAAGAC 742
QY 905 ATCTTGAAAGAAAGAGATGTTGTTCAATTAACAGAACAAAGAAAGATTAATCTGCTG 964
DB 743 ATCTTGAAAGAAAGAGATGTTGTTCAATTAACAGAACAAAGAAAGATTAATCTGCTG 802
QY 965 ACATCAAAATCATACATTAGAGACAGTAAATATTTCTGATGTTATGTTTCCAAATPA 1024
DB 803 ACATCAAAATCATACATTAGAGACAGTAAATATTTCTGATGTTATGTTTCCAAATPA 862
QY 1025 GTAAGTATGAGATGAATAGTGTCTGAATATGACGATTCAATTAATAGATGAAGAG 1084
DB 863 GTAAGTATGAGATGAATAGTGTCTGAATATGACGATTCAATTAATAGATGAAGAG 922
QY 1085 ATGATGAAGACTTAAACGAATTTAAGCCTATTTGCAATATGACAAATTTCCAAGTGAAG 1144
DB 923 ATGATGAAGACTTAAACGAATTTAAGCCTATTTGCAATATGACAAATTTCCAAGTGAAG 982
QY 1145 AAAACATAGGAATTTATTAAGAACTAGAGATTTGTATAGAGAAATGAAATTTAGATG 1204
DB 983 AAAACATAGGAATTTATTAAGAACTAGAGATTTGTATAGAGAAATGAAATTTAGATG 1042
QY 1205 ATTTAGATGAAGAAATAGAAATATCATCAGAGAAATTTATCTGAGAAAAATTAAGAAAG 1264
DB 1043 ATTTAGATGAAGAAATAGAAATATCATCAGAGAAATTTATCTGAGAAAAATTAAGAAAG 1102
QY 1265 GAAAGAAATATGAAAAACAAAGGATATATATTTTAAACCAATGATTAAGATTGTATG 1324
DB 1103 GAAAGAAATATGAAAAACAAAGGATATATATTTTAAACCAATGATTAAGATTGTATG 1162
QY 1325 ATGAGCATTTAAATAATTAATAATGATAGAGGTTTAAATAGAGAAAGAAATTCATCA 1384
DB 1163 ATGAGCATTTAAATAATTAATAATGATAGAGGTTTAAATAGAGAAAGAAATTCATCA 1222
QY 1385 TAAAAATCATTTGTTTCATATATTTTGAACGAGACATGAAATTTTACAGATCGTGATGAGT 1444

```

```

DB 1223 TAAATCATTTGTTTCATATATTTGACGAGCAATGAAATTTTACAGATCGTGATGAGT 1282
QY 1445 TATCTGAAGATATTAAGTAAATTTTATGAAACATTA 1482
DB 1283 TATCTGAAGATATTAAGTAAATTTTATGAAACATTA 1320

```

RESULT 4

ID AA005140 standard; DNA; 1566 BP.

XX AA005140;

XX AC 25-MAR-2003 (revised)

XX DT 05-NOV-1990 (first entry)

XX DE Sequence encoding N-terminal of peptide antigen to malarial sporozite.

XX KM Malaria; sporozite; vaccine; exoerythrocytic parasites; tetanus toxoid.

XX OS Plasmodium falciparum.

XX Key Location/Qualifiers

XX FT CDS 169..1554

XX FT /*tag= a

XX PN MO9006130-A.

XX PD 14-JUN-1990.

XX PF 30-NOV-1988; 88US-00278234.

XX PR 30-NOV-1988; 88US-00278234.

XX PR 12-APR-1989; 89US-00337204.

XX PA (BIOM-) BIOMEDICAL RES INST.

XX PI Hollingdal MR;

XX DR WPI; 1990-209624/27.

XX DR P-PsDB; AAR05766.

XX PT Novel malarial sporozite antigenic protein - useful as vaccine against sporozite(s) and exo-erythrocytic parasites.

XX PS Disclosure; Page 7; -pp; English.

XX CC Antigen, preferably linked at the C-terminal to a carrier such as tetanus toxoid, may be used as a vaccine against the malarial sporozite. (updated on 25-MAR-2003 to correct PI field.)

XX SQ Sequence 1566 BP; 745 A; 210 C; 330 G; 281 T; 0 U; 0 Other;

XX Query Match 39.6%; Score 586.4; DB 2; Length 1566;

XX Best Local Similarity 91.7%; Pred. No. 1.1e-74;

XX Matches 620; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 CAAGAACAAAGAGATCTTGAACAGAGAGAGCTGCTTAAGAAAGTTGCAAGACAA 60

DB 745 CAAGAACAAAGAGATCTTGAACAGAGAGAGCTGCTTAAGAAAGTTGCAAGACAA 804

QY 61 CAAGCGATTTAGAACAAAGATGAGCTGCTTAAGAAAGTTGCAAGAGCAGAT 120

DB 805 CAAGCGATTTAGAACAAAGATGAGCTGCTTAAGAAAGTTGCAAGAGCAGAT 864

QY 121 TTAAGACAGAGAGATCTTGAAGAAAGTTGCAAGACAAAGAGATCTAGACAA 180

DB 865 TTAAGACAGAGATCTTGAAGAAAGTTGCAAGACAAAGAGATCTAGACAA 924

QY 181 GAGAGCGTGTAAAGAAAGTTGCAAGACAAAGAGATCTTGAAGAAAGAGAGAGT 240

DB 925 GATAGAGCGTGTAAAGAAAGTTGCAAGACAAAGAGATCTTGAAGAAAGAGAGAGT 984

QY 241 GCTAAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAGTAGACTTGCTAAAGAA 300
 DB 985 GCTAAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAGTAGACTTGCTAAAGAA 1044
 QY 301 AAGTTACAAAGACGCAAGAAAGCGATTGAAACAAGTAGACTTGCTAAAGAAAGTTGCA 360
 DB 1045 ACGTTGCAAGACGCAAGAAAGCGATTGAAACAAGTAGACTTGCTAAAGAAAGTTGCA 1104
 QY 361 GAACCAACAAAGCGATTGAAACAAGACGCTGCTAAAGAAAGTTGCAAGAACAA 420
 DB 1105 GAACCAACAAAGCGATTGAAACAAGACGCTGCTAAAGAAAGTTGCAAGAACAA 1164
 QY 421 ACGATTGTAACAAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTG 480
 DB 1165 ACGATTGTAACAAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTG 1224
 QY 481 GAACCAAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTGTAACAAGAC 540
 DB 1225 GAACCAAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTGTAACAAGAC 1284
 QY 541 AAGCGCTGTAAGAAAGTTGCAAGAACAAAGCGATTGTAACAAGACGCTGCT 600
 DB 1285 AAGCGCTGTAAGAAAGTTGCAAGAACAAAGCGATTGTAACAAGACGCTGCT 1344
 QY 601 AAGAAAGGTTGCAAGACGCAAGAGATTTAGAACAAAGCGCTGATACGAAAAA 660
 DB 1345 AAGAAAGGTTGCAAGACGCAAGAGATTTAGAACAAAGCGCTGATACGAAAAA 1404
 QY 661 AATTGAAAGAAAA 676
 DB 1405 TTGCAAGAACAAAG 1420

RESULT 5
 ADO21941
 ID ADO21941 standard; DNA; 1371 BP.

AC ADO21941;

DT 12-AUG-2004 (first entry)

DE LSA-NRC(H) construct DNA derived from Malaria parasite LSA-1.

XX LSA-NRC, liver stage antigen; LSA-1; T-cell; B-cell epitope;

KM parasite sporozoite vacuole; antimalarial; vaccine; immunotherapy;

KW malaria parasite P. falciparum; LSA-NRC(H) construct; ds; gene.

XX Plasmidium falciparum.

OS Synthetic.

FT Key Location/Qualifiers
 FT CDS 1..1371
 FT /tag= a
 FT /product= "LSA-NRC(H) construct protein derived from
 FT Malaria parasite LSA-1"

FT WO200404167-A2.

XX 27-MAY-2004.

PD 12-NOV-2003; 2003WO-US036011.

PR 12-NOV-2002; 2002US-0425719P.

XX (REED-) REED ARMY INST RES WALTER.

PA Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
 PI Barbosa A;

XX WPI; 2004-420309/39.

DR P-PSDB; ADO21942.

XX

PT Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
 PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
 PT (LSA-1) epitope.
 XX
 PS Claim 16; SEQ ID NO 25; 90pp; English.
 XX
 CC The invention relates to a novel recombinant LSA-NRC polypeptide
 CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
 CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
 CC falciparum. The polypeptide of the invention demonstrates antimalarial
 CC activities and may be useful as a vaccine during immunotherapy and as a
 CC diagnostic agent, preferably for diagnosing malaria. The current sequence
 CC is that of the LSA-NRC(H)Mut construct DNA of the invention. The
 CC construct is harmonised for expression in Escherichia coli and comprises
 CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
 CC parasite P. falciparum LSA-1 in addition to a C-terminal His6 tag.
 XX

SQ Sequence 1371 BP; 539 A; 358 C; 283 G; 191 T; 0 U; 0 Other;
 Query Match 34.0%; Score 503.4; DB 12; Length 1371;
 Best Local Similarity 63.1%; Pred. No. 6; 9e-63;
 Matches 791; Conservative 0; Mismatches 461; Indels 1; Gaps 1;

QY 227 AACAGAGAGACGCTGTAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAGATA 286
 DB 79 AACAGAGAGACGCTGTAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAGATA 138
 QY 287 GACTTGCTAAAGAAAGTTGCAAGAGCAAGAAAGCGATTGAAACAAGAGACGCTGTA 346
 DB 139 AACAGAGAGACGCTGTAAGAAAGTTGCAAGAGCAAGAAAGCGATTGAAACAAG 198
 QY 347 AAGAAAGGTTGCAAGAACAAAGCGATTGAAACAAGAGACGCTGTAAGAAAGATA 405
 DB 199 AAGGATCCAGACCACTTTAAATCCCTCTGCGCAACCTCGGGGTTCCGAAACATC 258
 QY 406 TTGCAAGAACAAAGCGATTGTAAGAACAGACGCTGTAAGAAAGTTGCAAGATA 465
 DB 259 TTTCTCAAGAAAGAACTGAAACAGAGGCAACTGATTGAAACATTCATCAAGAC 318
 QY 466 CAACAAAGCGATTGTAAGAACAGAGACGCTGTAAGAAAGTTGCAAGAACAAAGC 525
 DB 319 GACGATGACAAAGAAAGTTGTAAGGCGCAGATGAAATCGCAGAGAACCTCGAA 378
 QY 526 GATTGTAAGAACAGAGACGCTGTAAGAAAGTTGCAAGAACAAAGCGATTGTAAG 585
 DB 379 GAAAGAGCTGCTGAACACAGCTGCACTGGAACGAGCGCTGCTAAAGAAAGCTC 438
 QY 586 CAAGAGACGCTGTAAGAAAGTTGCAAGAGCAAGCAAGAGATTGTAAGAACAAAG 645
 DB 439 CAGAGCGCTGCTGTAAGAAAGTTGCAAGAGCAAGCAAGCGACCTGGAACGCGCAAG 498
 QY 646 GCTGATAGCAAGAAAGTTGTAAGAAAGAAAGAAAGAGATATATATACAGAGAT 705
 DB 499 GCTGACAGAAAGAAAGTTGTAAGAAAGAAAGAAAGAGATATATATACAGAGAT 558
 QY 706 TTAATATGCTGTTGAAATATACAGCTATATAGACTTCAATGAAAGAAAGCTGAT 765
 DB 559 CTGTACGCGCTGTAAGAAATATACAGCTATATAGACTTCAATGAAAGAAAGCTGAT 618
 QY 766 TATATACCAATCAATCTTTTACCTGAGCAACAGAGGAAATATATAGATTCAG 825
 DB 619 TACATCCCAACCAAGAGACGCTGCAAGATATATATATATATATATATATATAT 678
 QY 826 GAAATATCTAT 885
 DB 679 GAAATCAAGATCAT 738
 QY 886 AAGGAT 945
 DB 739 CGCGACATCAACAAAGGCACTCGAAGAAAGAAAGAGCGCTCATCAACCAAGAAAG 798
 QY 946 AAGAGAGAT 1005

Qy	946	AAAGAAAGTAAATCTGCGACATACCAAAATCATATCATTAGAGACAGTAAATATTTCGAT	1005
Db	799	AAAGAAAGCAAAAGCGCTTGATATCCAGAACCAACCTCGTAGACCGTAAACATTAGCGAC	858
Qy	1006	GTTAATGATTTTCCAAATAGTAAGTAAATGATGAGATGAAATAAAGTGTGTAATATGACGATCA	1065
Db	859	GTGAACGACCTTCAGATCAGCAAGTACGAGACGAAATCTCCGCTGAAATAGATGATCC	918
Qy	1066	TTAAATAGATGAAGAAGAAATGATGAGAGACTTGAACGAAATTTAAGCCTATTGTGCAATAT	1125
Db	919	CTGATCGACGAAGAAGAAAGACGACGAATGATCGATGAATTTCAACCAATTTGTCCAGTAC	978
Qy	1126	GACAAATTTCCAAATGAGAAAGAAACATAGAAATTTATAAGAACTTGAAGATTTGATGAG	1185
Db	979	GATTAATTTCCAGACGAAGAAATATTCGGCAATTTACAAAGAACTCGAAGACCTCATCGAG	1038
Qy	1186	AAAAATGAAATTTTAGATGATTTTGTAGTGAAGAAATAGAAATATCATCGAAGATTAATCT	1245
Db	1039	AAAAACGAAACCTGAGCAGCTGAGCGAAGCAAGGCAATCGAAAAATCTCCGAAAGAACTGAGC	1098
Qy	1246	GAAGAAAAAATTAATAAAAGAAAGAAATATGAAAAAACAAGATTAATATTTTAAACCA	1305
Db	1099	GAAGAAAAAATCAAAAAAGCGCAAGAAATACGAAAAAACCAAGACACAACTTCAAAACA	1158
Qy	1306	AATGATAAAGATTGTATGATGAGCATTTAAAAAATTAATAAAATGATTAACAGATTAT	1365
Db	1159	AACGCAAAATCCCTCTACGACGAGACATTTAAAAAATACAAAAAGCAACCAAGTGAAC	1218
Qy	1366	AAGGAAAAAGAAAAATTCATATAATCATTTGTTCTATATATTTGACGAGACAAATGAATTT	1425
Db	1219	AAGGAAAAAGAAAAATTTATCAAAATCCCTCTCCACATCTTCGATGGCGAATTAAGAAATTT	1278
Qy	1426	TTACAGATCGGAGTGA---GTTATCGAAGATATTAATTAATATTTTATGAATCT	1478
Db	1279	CTGCAAATTTGAGCGAAACGGTTTGACGCAAGACATCACTAAATCTTCATCATAGACT	1334

RESULT	ID	AAQ28117	standard; DNA; 988 BP.
XX	AC	AAQ28117;	
XX	DT	25-MAR-2003 (revised)	
XX	DT	08-FEB-1993 (first entry)	
XX	DE	P.falciparum LSA gene 5' region.	
XX	KW	Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;	
XX	KW	paludism; liver stage-specific antigen; ss.	
XX	OS	Plasmodium falciparum.	
XX	FT	Key	Location/Qualifiers
XX	FT	5' UTR	1..32
XX	FT	misc_feature	/*lag= a
XX	FT		33..988
XX	FT		/*lag= d
XX	FT	repeat_region	/note= "5' part of LSA gene"
XX	FT		492..988
XX	FT		/*lag= b
XX	FT	repeat_unit	492..542
XX	FT		/*lag= c
XX	PN	MO9213884-A1.	
XX	PD	20-AUG-1992.	
XX	PP	05-FEB-1992;	92MO-FR000104.
XX	PR	05-FEB-1991;	91FR-00001286.
XX			

PA	(INSP) INST PASTEUR.
XX	
XX	Guérinmarchand C, Drulhe P;
XX	WPI; 1992-299985/36.
DR	P-PSDB; AAR26943.
XX	
PT	Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT	vaccination against, treatment of and diagnosis of malaria.
XX	
PS	Claim 11; Fig 6; 81pp; French.
XX	
CC	This nucleotide sequence is the 5' part of the P.falciparum liver-stage
CC	specific antigen (USA) gene. It codes for a polypeptide sequence which
CC	carries a T cell epitope characteristic of a protein produced in
CC	hepatocytes infected with P.falciparum. The polypeptide can be used in
CC	the preparation of vaccines against malaria. (Updated on 25-MAR-2003 to
CC	correct PN field.)
XX	
Q0	Sequence 988 BP; 465 A; 120 C; 184 G; 219 T; 0 U; 0 Other;

Query Match	31.9%;	Score 472.8;	DB 2;	Length 988;
Best Local Similarity	95.7%;	Pred. No. 1.5e-58;		
Matches 486;	Conservative 0;	Mismatches 22;	Indels 0;	Gaps 0;

Qy	77	AAAGTAACTTGGCTAAAGAAAAGTAAACAAGGCGCAAAAGCAATTTTGAACAAGAGAGAC	136
Db	481	AAAGAAAAAGCACTAAAGAAAAGTTTACAGGGGCAACAAACGATTTCAAGACAGAGAGAC	540
Qy	137	TTGCTAAAGAAAAGTTGCTCAAGAACAAACAAGCGACTAGAACAAAGAGACGTCCTAAAG	196
Db	541	GTGCTAAAGAAAAGTTGCTCAAGAACAAACAAGCGATTTAGAACAAAGAGACCTTGTAAAG	600
Qy	197	AAAAGTTGCAAGAACAAACAAGCGATTTAGAACAAAGAGACGTCCTAAAGAAAAGTTGC	256
Db	601	AAAAGTTGCAAGAACAAACAAGCGATTTAGAACAAAGAGACGTCCTAAAGAAAAGTTGC	660
Qy	257	AAAGAACAAACAAGCGATTTTGAACAAGATGACTTGTCTAAAGAAAAGTTACAAAGACGAC	316
Db	661	AAAGAACAAACAAGCGATTTTGAACAAGAGACCTTGTCTAAAGAAAAGTTTCAAGAACAC	720
Qy	317	AAAGCGATTTTGAACAAGAGACGTCCTAAAGAAAAGTTGCAAGAACAAACAAGCGATT	376
Db	721	AAAGCGATTTTGAACAAGAGACGTCCTAAAGAAAAGTTGCAAGAACAAACAAGCGATT	780
Qy	377	TAGAACAAAGAGACGTCCTAAAGAAAAGTTGCAAGAACAAACAAGCGATTTAGAACAG	436
Db	781	TAGAACAAAGAGACGTCCTAAAGAAAAGTTGCAAGAACAAACAAGCGATTTAGAACAG	840
Qy	437	AGAGACTTGTAAAGAAAAGTTGCAAGAACAAACAAGCGATTTAGAACAAAGAGACGTC	496
Db	841	AGAGACTTGTGTAAAGAAAAGTTCAAGAGACGCAAAAGCGATTTAGAACAAAGATGACTTG	900
Qy	497	CTAAAGAAAAGTTGCAAGAACAAACAAGCGATTTAGAACAAAGAGACGTCCTAAAGAAA	556
Db	901	CTAAAGAAAAGTTGCAAGAACAAACAAGCGATTTAGAACAAAGAGACGTCCTAAAGAAA	960
Qy	557	AGTTGCAAGAACAAACAAGCGATTTTGA	584
Db	961	GGTTGCAAGAACAAACAAGCGATTTTGA	988

ADP85917	RESULT 8
ID	ADP85917 standard; DNA; 1300 BP.
XX	
AC	ADP85917;
XX	
DT	26-AUG-2004 (first entry)
XX	
DE	Synthetic construct #1.
XX	
CM	Nanopore data analysis ; polymer ; polynucleotide;

KM single-molecule resolution; ds.
 XX
 OS Synthetic.
 XX
 PN US2004110205-A1.
 XX
 PD 10-JUN-2004.
 XX
 PF 23-SEP-2003; 2003US-00668749.
 XX
 PR 23-SEP-2002; 2002US-0412959P.
 XX
 PA (WANG/) WANG H.
 XX
 PI Wang H;
 XX
 DR WPI; 2004-467656/44.
 XX
 PT Nanopore analysis system for analyzing polynucleotides includes nanopore
 PT device having structure with aperture, and nanopore data analysis system
 PT that can generate nanopore data points corresponding to each target and
 PT non-target polymer.
 XX
 BS Disclosure; SEQ ID NO 1; 17pp; English.

The invention relates to a nanopore analysis system that comprises a nanopore device and a nanopore data analysis system. The nanopore device has a structure having an aperture. The nanopore data analysis system can generate nanopore data points corresponding to each target polymer and each non-target polymer traversing the aperture, form a distribution pattern of the data points, and analyse a distribution of target polymer data points in the distribution pattern. The distribution pattern includes data cluster(s). The nanopore data analysis system is stored on a computer-readable medium. A mechanism is provided for analysing the distribution of target polynucleotide data points in the distribution pattern. The nanopore data analysis system can also analyse the distribution of the non-target polynucleotide data points. It can also determine a ratio between the target polynucleotide data points and the non-target polynucleotide data points. It can further analyse the distribution of target polynucleotide data points between the two data clusters; compare the distribution of the target polynucleotide data points between the two data clusters to a phosphorylation state standard distribution; and determine a ratio of phosphorylated target polynucleotide to non-phosphorylated target polynucleotides. It can also determine a cluster score for the target polynucleotide data points in a defined area, and compare the cluster score for the target polynucleotide data points to a cluster score for a chemical integrity standard density distribution for the defined area in a distribution of a target polynucleotide standard. The invention is useful for analysing polymers, such as polynucleotides. The inventive nanopore analysis system potentially provides high speed sampling with single-molecule resolution, which may enable unprecedented dynamic range and sensitivity in analysis of samples containing charged polymers such as polynucleotides or polypeptides. The present sequence is a synthetic construct used in the invention.

	Query Match	Similarity	Score	DB	Length	Mismatches	Gaps
Qy	91	AAAGAAAATTACAGAGCAGCGATTTAGAACACAGAGACTTGCTTAAGAAAG	15.7%	233.4	1300	48.7%	Pred. No. 1.5e-24;
Dd	1	AAA	633	Conservative	0	Mismatches	666; Indels 0; Gaps 0.
Qy	151	TTCGAGAACAACAAGCGATTAGAACAGAGAGACGTCTAAAGAAAGTTGCAGAA					
Dd	61	AAA					
Qy	211	CACCAAGCGATTTAGAACACAGAGACGTCTAAAGAAAGTTCAGACACAAAGC					
Dd	121	AAA					

[illegible][illegible]


```

Db      2933 ATGAAGAAAGAAAGAAACACATGAAGAAAGAAATGTAACTATGAAAGAAAG 2992
Qy      1319 TGTATGATGACATATTAAAAATTAATAATGATTAACAGATTAAATAGAAAAAGAAA 1378
Db      2993 AAGAAAAAGTACACATGAAGAAAGAAAGTAAACATGAAGAAAGAAAGAAAGTAA 3052
Qy      1379 AATTCAATAAATTCATTGTTTCATATATTGACGAGACAAATGAAATTTTACAGATCGTG 1438
Db      3053 CACATGAAGAAAGAAAGAAATGTAACTATGAAAGAAAGAAAGAAAGTAACTACATGAAG 3112
Qy      1439 ATGAGTTATCTGAAGATTAACCTAA 1463
Db      3113 AAGAAAGAAAGAAAGTAAATGAAGAA 3137

RESULT 10
AAN80899
ID      AAN80899 standard; DNA; 210 BP.
XX
AC      AAN80899;
XX
DT      14-MAY-2003 (revised)
DT      25-MAR-2003 (revised)
DT      19-SEP-1990 (first entry)
XX
DE      Sequence encoding an epitope characteristic of a protein produced in
DE      liver cells infected with Plasmodium falciparum.
XX
KW      Plasmodium falciparum; vaccine; malaria diagnosis;
KW      malaria antibody detection.
XX
OS      Plasmodium falciparum.
XX
FH      Key      Location/Qualifiers
FT      CDS      1..204
FT      /*tag= a
XX
PN      WO8805785-A.
XX
PD      11-AUG-1988.
XX
PF      09-FEB-1988; 88WO-FR000074.
XX
PR      09-FEB-1987; 87FR-00001543.
XX
PA      (INSP ) INST PASTEUR.
PA      (MARC/) MARCHAND C.
PA      (CNRS ) CENT NAT RECH SCI.
XX
PI      Drulibe P, Pujjalome O, Marchand C, Langsley G;
XX
DR      WPI; 1988-235148/33.
DR      P-PSDB; AAP81355.
XX
PT      New polypeptide(s) contg. plasmodium falciparum epitope - useful for
PT      malaria vaccination or diagnosis.
XX
PS      Example 1; Fig 1; 37pp; French.
XX
CC      This sequence is contained in clone DG307 of a genomic DNA library of P.
CC      falciparum. AAN80899 is composed entirely of a repeating motif of 51 base
CC      pairs. The type of replication coded by DG307 seems to be conserved which
CC      implies that it may be present in all P. falciparum strains. The AA
CC      repeat unit is the basis of the peptides claimed in the patent, which are
CC      useful for the prodn. of of malaria vaccines and as immunosassay reagents
CC      for diagnosis of malaria by detection of antibodies. (Updated on 25-MAR-
CC      2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
CC      (Updated on 14-MAY-2003 to correct PS field.)
XX
SQ      Sequence 210 BP; 97 A; 33 C; 47 G; 33 T; 0 U; 0 Other;
Query Match      12.4%; Score 183.2; DB 1; Length 210;

```

```

      Best Local Similarity 95.9%; Pred. No. 2.2e-17;
      Matches 188; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      83 GACTTCTTAAAGAAAGTACTACAGACGAAACGATTTTGAACAGAGACCTTGGTA 142
Db      8 GACTTGTCTTAAAGAAAGTTCACAGACGAAACGATTTTGAACAGAGACCTTGGTA 67
Qy      143 AAGAAAGTTGCAAGAACCAAGCGATCTTGAACAGAGACGCTTAAAGAAAGT 202
Db      68 AAGAAAGTTGCAAGAACCAAGCGATCTTGAACAGAGACCTTAAAGAAAGT 127
Qy      203 TGCAGAACCAAGCGATTTAGAACAGAGACGCTTAAAGAAAGTTCAGAAC 262
Db      128 TACAGGGGCAACAAAGCGATCTTGAACAGAGACCTTGAAGAAAGTTCAGAAC 187
Qy      263 AACAAAGCGATTTTGA 278
Db      188 AACAAAGCGATTTGA 203

RESULT 11
ADQ62833
ID      ADQ62833 standard; RNA; 1000 BP.
XX
AC      ADQ62833;
XX
DT      07-OCT-2004 (first entry)
XX
DE      Homopoly-A contaminant for RNaseH activity assay.
XX
KW      ss; nuclease-mediated cleavage; target nucleic acid; RNaseH activity;
KW      fluorophore; fluorescence quencher; fluorescent signal; HIV;
KW      reverse transcriptase.
XX
OS      Synthetic.
XX
PN      WO2004059012-A1.
XX
PD      15-JUL-2004.
XX
PF      22-DEC-2003; 2003WO-US040879.
XX
PR      23-DEC-2002; 2002US-0436125P.
XX
PA      (AMHP ) WYETH.
XX
PI      Olson MW, O'Connell JF;
XX
DR      WPI; 2004-543471/52.
XX
PT      Detecting a nuclease-mediated cleavage of a target nucleic acid, useful
PT      for detecting and monitoring RNase H activity, comprises hybridizing a
PT      target nucleic acid to a fluorescently labeled oligonucleotide probe.
XX
PS      Example 1; SEQ ID NO 4; 61pp; English.
XX
CC      The invention relates to a method of detecting a nuclease-mediated
CC      cleavage of a target nucleic acid or measuring a RNaseH activity of an
CC      agent by hybridizing a target nucleic acid to a fluorescently labelled
CC      oligonucleotide probe complementary to the target nucleic acid and
CC      containing a fluorophore at one terminus and a quenching group at the
CC      other terminus and contacting the probe-target hybrid with an agent
CC      having nuclease activity. When the oligonucleotide probe is unhybridised
CC      to the target nucleic acid, the probe adopts a conformation that places
CC      the fluorophore and quencher in such proximity that the quencher quenches
CC      the fluorescent signal of the fluorophore. Formation of the probe-target
CC      hybrid causes sufficient separation of the fluorophore and quencher to
CC      reduce quenching of the fluorescent signal of the fluorophore followed by
CC      contacting the probe-target hybrid with an agent having nuclease activity
CC      to selectively cleave the target nucleic acid and thus release the intact
CC      probe. Detection of the release of the probe is by measuring a decrease
CC      in the fluorescent signal of the fluorophore as compared to the signal of
CC      the probe-target hybrid. The methods are useful for detecting and

```


CC monitoring RNaseH activity in real time. The methods are also useful for
 CC monitoring enzymatic degradation of an RNA-DNA duplex. In an example of
 CC the invention, the activity of the RNaseH from E. coli or HIV reverse
 CC transcriptase is measure by the method of the invention. In order to test
 CC the specificity of the RNaseH cleavage of the target-probe complex, a DNA
 CC or RNA contaminant is added to the assay reaction. This sequence
 CC corresponds to the single stranded homopolymeric polyA RNA contaminant
 CC for testing this activity.

XX Sequence 1000 BP, 1000 A; 0 C; 0 G; 0 T; 0 U; 0 Other;

Query Match 12.2%; Score 180.2; DB 12; Length 1000;

Best Local Similarity 48.8%; Pred. No. 5.3e-17;

Matches 405; Conservative 0; Mismatches 508; Indels 0; Gaps 0;

```

QY 397 AAAAAAAAAAGTTGACAGACAAACGCGATTGAAACAAGAGAGACTTGCTAAAGAAAAG 456
   |||||
DB 1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60
   |||||
QY 457 TTGCAAGAACAAACGCGATTGAAACAAGAGAGCGCTTAAAGAAAAGTTGCAAGAA 516
   |||||
DB 61 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 120
   |||||
QY 517 CAACAAGCGATTGACAGACGCGCTTAAAGAAAAGTTGCAAGAACAAACG 576
   |||||
DB 121 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 180
   |||||
QY 577 GATTGAAACAAGAGAGCGCTGTAAGAAAAGTTGCAAGACGCAAGAAAGATTGAA 636
   |||||
DB 181 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 240
   |||||
QY 637 CAAGAAGCGCTGATCGAATAAAATTGAAAGAAAAGCAATGAGATATATTA 696
   |||||
DB 241 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 300
   |||||
QY 697 GCAGAGATTTATATGCTGTTAGAAATACGAGATATGAACTTCATCAGAAATGAA 756
   |||||
DB 301 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 360
   |||||
QY 757 CGTGATATTATATACCATCATCTCTTCTTCTCAGACAAACGAGGAATAGTA 816
   |||||
DB 361 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 420
   |||||
QY 817 GATTCAAGGAATATCTATATAGAAAACAAATAGAGATCTATTCAACAATGTT 876
   |||||
DB 421 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 480
   |||||
QY 877 GAAGACGAAGGATATATACATTAAGGACATCTTGAAGAAAAGAGTTCATTA 936
   |||||
DB 481 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 540
   |||||
QY 937 CGAAGAACAAAGAGATTAATCTGCTGACATACAAATCATATTAGACAGTAA 996
   |||||
DB 541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 600
   |||||
QY 997 ATTCTGATGTTAATGATTTTCAATTAAGTATGAGAGTAAATTAAGTCTGAT 1056
   |||||
DB 601 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 660
   |||||
QY 1057 GACGATTCATTAATGATGAGAGAGATGATGAAGTCTTGAAGAACTTTAAGCT 1116
   |||||
DB 661 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 720
   |||||
QY 1117 GTGCAATATGACAAATTTCCAGATGAGAAAACATAGAAATTTATAGAGTAA 1176
   |||||
DB 721 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 780
   |||||
QY 1177 TTGATAGAGAAAATTTGATGATTTTATGATGAAAGTAAATCATCAGAA 1236
   |||||
DB 781 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 840
   |||||
QY 1237 GAATTAATCTGAGAAAATTAAGAAAAGAAATATGAAAACAAAGGATATAT 1296
   |||||

```

```

DB 841 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 900
QY 1297 TTTAAACCAATGATTAATGTTGATGATGACATATTTAAATAATATGATTAAG 1356
   |||||
DB 901 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 960
   |||||
QY 1357 CAGGTTAATAGAAAAGAAAATTCATTA 1389
   |||||
DB 961 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 993
   |||||

RESULT 12
AD062832/c
ID AD062832 standard; RNA; 1000 BP.
XX
AC AD062832;
XX
DT 07-OCT-2004 (first entry)
XX
DE Homopoly-U contaminant for RNaseH activity assay.
XX
KW ss; nuclease-mediated cleavage; target nucleic acid; RNaseH activity;
KM fluorophore; fluorescence quencher; fluorescent signal; HIV;
KM reverse transcriptase.
XX
OS Synthetic.
XX
PN WO2004059012-A1.
XX
PD 15-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US040879.
XX
PR 23-DEC-2002; 2002US-0436125P.
XX
PA (AMDP ) WYETH.
XX
PI Olson MM, O'Connell JF;
XX
DR WPI; 2004-543471/52.
XX
PT Detecting a nuclease-mediated cleavage of a target nucleic acid, useful
PT for detecting and monitoring RNase H activity, comprises hybridizing a
XX target nucleic acid to a fluorescently labeled oligonucleotide probe.
XX
PS Example 1; SEQ ID NO 3; 61pp; English.
XX
XX
CC The invention relates to a method of detecting a nuclease-mediated
CC cleavage of a target nucleic acid or measuring a RNaseH activity of an
CC agent by hybridizing a target nucleic acid to a fluorescently labelled
CC oligonucleotide probe complementary to the target nucleic acid and
CC containing a fluorophore at one terminus and a quenching group at the
CC other terminus and contacting the probe-target hybrid with an agent
CC having nuclease activity. When the oligonucleotide probe is unhybridised
CC to the target nucleic acid, the probe adopts a conformation that places
CC the fluorophore and quencher in such proximity that the quencher quenches
CC the fluorescent signal of the fluorophore. Formation of the probe-target
CC hybrid causes sufficient separation of the fluorophore and quencher to
CC reduce quenching of the fluorescent signal of the fluorophore followed by
CC contacting the probe-target hybrid with an agent having nuclease activity
CC to selectively cleave the target nucleic acid and thus release the intact
CC probe. Detection of the release of the probe is by measuring a decrease
CC in the fluorescent signal of the fluorophore as compared to the signal of
CC the probe-target hybrid. The methods are useful for detecting and
CC monitoring RNaseH activity in real time. The methods are also useful for
CC the invention, the activity of the RNaseH from E. coli or HIV reverse
CC transcriptase is measure by the method of the invention. In order to test
CC the specificity of the RNaseH cleavage of the target-probe complex, a DNA
CC or RNA contaminant is added to the assay reaction. This sequence
CC corresponds to the single stranded homopolymeric polyU RNA contaminant
CC for testing this activity.
XX

```

SQ Sequence 1000 BP; 0 A; 0 C; 0 G; 0 T; 1000 U; 0 Other;

Query Match 12.2%; Score 180.2; DB 12; Length 1000;
Best Local Similarity 48.8%; Pred. No. 5.3e-17;
Matches 485; Conservative 0; Mismatches 508; Indels 0; Gaps 0;

```

Qy 397 AAAGAAAGTTGCAAGAACAAACAAAGCGATTAGAACAGAGAGACTTGTAAGAAAG 456
Db 1000 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 941
Qy 457 TTGCAAGAACAAACGAGTTAGAACAGAGAGAGCTGCTAAAGAAAGTTGCAAGAA 516
Db 940 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 881
Qy 517 CAACAAAGCGATTGAAACAAGAGAGAGCTGCTAAAGAAAGTTGCAAGAACAAAGC 576
Db 880 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 821
Qy 577 GATTAGAACAAAGAGAGAGCTGCTAAAGAAAGTTGCAAGAGAGAGATTAGAA 636
Db 820 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 761
Qy 637 CAAGAGAGAGCTGATCGAAGAAATTTAGAGAGAGAGAGAGAGAGAGAGATATTA 696
Db 760 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 701
Qy 697 GCAGAGATTATATATGCTGCTTAAAGAAATACAGCTATAGAACTTCCATGAAAGTAA 756
Db 700 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 641
Qy 757 CGTGATTTATATACACATCAATCTTTACTCTGAGACAAAGAGAGAGATAGTAA 816
Db 640 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 581
Qy 817 GATTCCAGGAATCTATATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 876
Db 580 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 521
Qy 877 GAAGAGAGAGAGATATATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 936
Db 520 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 461
Qy 937 CCAGAGAGAGAGAGATATATCTGCTGACATACATACATACATACATACATACAT 996
Db 460 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 401
Qy 997 ATTCTGATGTTATGATTTTCAATTAATTAATTAATTAATTAATTAATTAATTA 1056
Db 400 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 341
Qy 1057 GACGATTGATTATGATGAAGAGAGATGATGAAGAGCTTGAAGAGATTGAAGCTATT 1116
Db 340 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 281
Qy 1117 GTGCAATATGCAATTTCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1176
Db 280 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 221
Qy 1177 TTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1236
Db 220 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 161
Qy 1237 GAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTAT 1296
Db 160 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 101
Qy 1297 TTTAAACCAATGATTAAGTTTGTATGATGAGAGAGAGAGAGAGAGAGAGAGAG 1356
Db 100 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 41
Qy 1357 CAGGTTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1389
Db 40 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8

```

RESULT 13
AA087587
ID AA087587 standard; DNA; 1686 BP.

AA087587;

27-AUG-2003 (revised)
19-DEC-1995 (first entry)

DNA encoding Leucocytozoan protozoa structural protein epitope.

leucocytozoan protozoa; structural protein; epitope; vaccine; fowl;

leucocytozoanosis; treatment; ss.

Leucocytozoan.

JP07089995-A.

04-APR-1995.

10-SEP-1993; 93JP-00226078.

10-SEP-1993; 93JP-00226078.

(DOBU) DOBUTSUO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.

(NIRS-) NISSEIKEN KK.

WI; 1995-167252/22.

P-PSDB; AAR70491.

Immune inducing polypeptide against Leucocytozoan protozoa - useful in production of vaccines for treatment of Leucocytozoanosis in fowl.

Claim 1; Page 12-14; 20pp; Japanese.

AA087587-89 encode polypeptides having a whole or partial epitope of a structural protein of Leucocytozoan protozoa (see AAR70491-93). The polypeptides and DNA encoding them are useful in the production of vaccines for the treatment of Leucocytozoanosis of fowl. (Updated on 27-AUG-2003 to correct OS field.)

Sequence 1686 BP; 915 A; 129 C; 382 G; 260 T; 0 U; 0 Other;

Query Match 12.1%; Score 179.4; DB 2; Length 1686;
Best Local Similarity 46.9%; Pred. No. 6.6e-17;
Matches 595; Conservative 0; Mismatches 671; Indels 3; Gaps 1;

```

Qy 2 AAGAACAAACGAGATCTAGAACAGAGAGAGCTGCTAAAGAGAGAGAGAGAGAG 61
Db 101 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 160
Qy 62 AAGCGATTAGAACAGATGACTTCTAAAGAAAGTTACAGAGAGAGAGAGAGATT 121
Db 161 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 220
Qy 122 TAGAACAGAGAGAGCTTCTAAAGAAAGTTGCAAGAGAGAGAGAGAGAGAGT 181
Db 221 ATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 280
Qy 182 AGAGAGCTGCTAAAGAAAGTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 241
Db 281 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 340
Qy 242 CTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
Db 341 ATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 400
Qy 302 AGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
Db 401 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 460

```

Oy		362	AACCAAAAGCCGTTTAAACAAGAAGACGTCTAAAGAAAAGTTGCCAAGAACCAAA	421
Db		461	AAGAACCAACAGAACACAGATGAAGAAAGAACCAAGATGAAGAAACAAGATGAAGAAAG	520
Oy		422	GCGATTTCAGACAGAGACTTGCTTAAGAAAAAGTTGCAGAGAACCAAGCGATTAG	481
Db		521	AAGAAAGAAACACGAGAAACAGATGAAAGAACAGAGAAAGATATGCTGAAAAAGAAA	580
Oy		482	AACAGAGAGACGTGCTTAAGAAAAAGTTGCAGAGAACCAAGCGATTAGAACAGAGA	541
Db		581	ATGAAGATGAAAGAAAAAGAAAAAGAAAGAAACAGAGAGATGAAGAAAAATATATGTTG	640
Oy		542	GACGTGCTTAAGAAAAAGTTGCCAAGAACCAAGCGCTTTTAGAACAGAGAGACGTGCTA	601
Db		641	AAAAAGAAAAAGATGAAGATATATGTTTTGAAAAAGAACCAAGAACATTTTAAAGAAATTAG	700
Oy		602	AAGAAAAAGTTGCAGAGACAGCAAAAGAGATTTAGAACAAAGAGAAAGCTGATACGAAAAAA	661
Db		701	AAAAAGAACTTTTAGAAGAAAAAGTGAAGAAAGTAAACA---TAAAGATGATTAAGAAAGACA	757
Oy		662	ATTAGAAAAAGAAAAAGAACATGAGATATATTAGCAGAGATTTATATGTCGTTTAG	721
Db		758	CAAAAATMAAAAGCATGAAGTTGAGAGAGAAACGAGAAATAGAAATTAAGAAAGGAATCAG	817
Oy		722	AAATACCAGCTTAGAACCTTCCATCAAGAAATGAACGTGGATTTATATATCCATCAAT	781
Db		818	AAGAAATATGATAGAAAAAAATCAAGAGTAAAGAAACAAATGAGCAAGTCAAGAGAAAAA	877
Oy		782	CTTCTTTTCCGAGGCAACAGAGGAAATAGTAGAATTCAGAAATTCATATAATAG	841
Db		878	CAATGAAACAAAAAGAAATATATACAGATCCAAATTTGGTCATCAATTTAAGATATATGA	937
Oy		842	AAAAAACCAATAGAGAAATCTATTACAACAAATGTTGAAGACAGAGGATATATCATAAAG	901
Db		938	TAAATTAAGAAACAGAAACAGAAAAACATATCAGCAAATTTATACATTAAAGTAACTA	997
Oy		902	GACATCTTGAAGAAAAAGAAAGATGTTCAATMAAACCAAGAACAAAAAGAAAGATTAATCTG	961
Db		998	ATGAATCAAAATCCGAAATMAAATTCACATCAAAATATATTTGAAGATGAAGAAATPAAAAATT	105
Oy		962	CTGACATACAAATTCATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTCCAA	102
Db		1058	CATCATTTGAAAAAACCGAAATTTAAATGAAGCTACTAGAGGCCGGAATTCGGAACAATAG	111
Oy		1022	TAACTAAGATGAGAGATGAATTAAGTCTGAATATGACGATTCATTAATATGATGAAGAG	108
Db		1118	TTGAAGAAAGTTTGTGAAGAAACCCAGTATGTTGAGAGAAAGAAATGTAAGAGAAACCTG	117
Oy		1082	AAGATGATGAAGACTTAGACGATTTTAAGCTATATGTCAAATATGACAAATTTCCAAGATG	114
Db		1178	TAGTTGAAGAAAGAAATGATGAAGAAACACCTGATTTGAAGAAAGAAATGATGAAGAAA	123
Oy		1142	AAGAAAAACATAGGAATTTTAAAGAACTGAAGATTTGATAGAGAAAAATGAATAATTAG	120
Db		1238	CACCACTAGTGAAGAGATGTCGTGCAAGAAACACCAATAGTGGAAAGAAAGATGATGAG	129
Oy		1202	ATGATTTAGTAGAAGATGAAGAAAAATCATCGAAGAAATTAATCTGAAGAAAAATTAATA	126
Db		1288	AAGAAACTAATATTAATTTGAAGAAAGTATGAAGAAACACCAATGATTTGAAGAAAGAAATGA	135
Oy		1262	AAGGAAGA 1270	
Db		1358	TAGAGAAA 1366	
RESULT 14				
AAA70099 standard; DNA; 3579 BP.				
XX	AAA70099;			
AC				
XT	07-NOV-2000 (first entry)			

DE	Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:232.
KW	Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
XV	antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX	
OS	Plasmodium falciparum.
XX	
PN	WO200025728-A2.
PD	
XX	
PP	11-MAY-2000.
PF	05-NOV-1999; 99WO-US026796.
PR	05-NOV-1998; 98US-0107131P.
PA	(HOFF/) HOFFMAN S.
PA	(CARU/) CARUCCI D.
PA	(GARD/) GARDNER M.
PA	(VENT/) VENTER J C.
PI	Hoffman S, Carucci D, Gardner M, Venter JC;
DR	WPI: 2000-365347/31.
PT	Proteins encoded by chromosome 2 of the human malarial parasite,
PT	Plasmodium falciparum, useful as antimalarial vaccines and in the
XX	diagnosis of P.falciparum infection.
PS	Disclosure: Page 457-458; 577pp; English.
XX	
CC	The present invention describes proteins and their fragments (I) encoded
CC	by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC	Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
CC	vaccines against P. falciparum infection comprising (I) or (II). (I) and
CC	(II) are useful for the development of vaccines against P. falciparum
CC	infection. (I) and polyclonal antisera or a monoclonal antibody raised to
CC	immunogens comprising the sequences of (I), are useful in the detection
CC	of infection with P. falciparum. Furthermore, (I) (especially when they
CC	are rifins or secreted or membrane proteins) can aid the identification
CC	of drugs to treat or prevent P. falciparum infection, or they can be used
CC	to identify drug resistance in P. falciparum. Sequencing of the
CC	Plasmodium chromosome 2 and the subsequent identification of proteins
CC	encoded by it will help to expand our understanding of parasite biology,
CC	a process hampered by the complexity of the parasitic lifecycle, and
CC	provide new targets for vaccine and drug development. Parasite resistance
CC	to drugs and mosquito resistance to insecticides have led to a resurgence
CC	of malaria in many parts of the world, and there is a pressing need for
CC	vaccines and new drugs. AA#70078 to AA#70287 and AAB18144 to AAB18352
CC	represent nucleotide and protein sequences given in the present
CC	invention, but which are not specifically mentioned within the
CC	specification
XX	
SEQ	Sequence 3579 BP; 1904 A; 398 C; 552 G; 725 T; 0 U; 0 Other;
Query Match	11.9%; Score 176.6; DB 3; Length 3579;
Best Local Similarity	46.3%; Pred. NO. 1.6e-16;
Matches	661; Conservative 0; Mismatches 755; Indels 9; Gaps 2
OY	36 TGCTAAGAAAGAAGTTGCAGAACCAACAAGCGATTTAGACAGACTTGCTTAAGA 95
DB	744 TGATTAATTAACAAAGAAAAAAAATACTACAAACACTGTGAAAAAAACAAAAAGAACAGAGA 803
OY	96 AAAAGTTCACAGAGCAGCAAAGCGATTTCAGAACAGAGAGACTTGTAAGAAAGTTGCA 155
DB	804 TAAGAAACATTCTAAAAAAAAGAAAAACAAAACAAATTTTAATTAAGGAAGAAAAACCA 863
OY	156 AGAACCAACAAAGCATCTAGAACCAAGAGAGACGTCTAAAGAAAAAGTTGCAGAACACACA 215
DB	864 AAAATATGAAAAATCATATATCTCCAAGTAATTTAAACAGACAGCAATGGAAATTTTGA 923
OY	216 AAGCGATTTAGACAGAGAGACGTGCTAAAGAAAGTTGCAGAACCAACAAAGCATTT 275
DB	924 TGATTAATGAAAAAGTAATGGAATCCGACAAAAATTTTAAAGAGAAAAAGGGGGAAATGA 983

Qy	27	AGAA	CBAATG	ACTTGCTG	TAAAGAAA	GGTTTCA	AAGACG	CAACCG	CAATTTG	AACA	GA	335								
Db	984	AAAA	CTGATCT	TAATAT	ATBGCAT	TTATATA	AAAAA	GAAAGAA	GGAACAAA			1043								
Qy	336	GAGC	GTGCT	TAAGAAA	GGTTC	CAAGAA	CAACAA	AGCGAT	TTAGAAC	CAAGAG	CGTGC	395								
Db	1044	AGAT	GA	AAAAAG	AAAAA	CATTCA	CAGCTAG	TAAAG	TGTAC	ATTAA	TTAAAG	1109								
Qy	396	TAA	AAAAAG	TTG	CAAGAA	CAACAA	AGCGAT	TTTAGA	CAAGAG	AC	CTTGT	AAAA	455							
Db	1104	AGAA	AAAA	CGAAAG	CGACG	AAAAA	GAAGAC	CAAAAG	GAACAG	ATGAT	TGAAG	GAATAC	1163							
Qy	456	GTTG	CAAGAA	CAACAA	AGCGAT	TTTAGA	CAAGAG	ACG	TGCTTAA	GAAGAA	GGTTC	CAAG	515							
Db	1164	AGAT	GTATGA	AGATG	ATACAG	ATGAT	GAAGAG	ATACAG	ATGATGA	AGATG	ATCA	AGTGA	1222							
Qy	516	ACAA	CAAA	-----	CGCAT	TTTAGA	CAAGAG	ACG	TGCTTAA	GAAGAA	GGTTC	CAAG	569							
Db	1224	TGAA	AAAA	CTACAG	GTGAT	CAAGAAAA	CAAGAA	GAACAG	ATGATG	ACG	AAAAA	AAAAA	1283							
Qy	570	ACAA	AGCAT	TTTAGA	CAAGAG	ACG	TGCTTAA	GAAGAA	GGTTC	CAAGAG	CGCA	AGAA	629							
Db	1284	AGAA	AAAC	CCGAG	AGAGAT	TTTAGA	GAAGAC	CAAAAG	AGATG	AGAAAA	GAACAA	GAAGA	1343							
Qy	630	TTT	AGAA	CAAGAA	AGCGTGA	TACG	AAAAA	AAAAAT	TAGAA	GA	AAAAA	AGAACAT	GAAG	689						
Db	1344	ATCA	AAAAA	AGATTA	AGAAAG	ATCAG	AAAAA	AGATTA	AGAAAG	ATCAG	AAAAA	AGACAA	GA	1403						
Qy	690	TAT	ATTAG	ACAGAT	TTTATAT	ATG	CGTTTAG	AAAT	TACAG	CTATAG	AACTT	CCAT	CA	749						
Db	1404	AAAA	CTAAG	AAAGAT	TGA	AAAAA	CTGA	AGACG	AAAAA	GGAGCG	AGAT	TAC	AAAA	1463						
Qy	750	AAAT	GAA	CGTGA	ATTATAT	TATCA	CATCA	ATCTT	TATCCT	CAAG	CAACAG	AGAG	GAA	809						
Db	1464	AGAA	AC	ATG	ATG	ATG	AAAAA	AAAGG	AAAT	TGAG	AGAG	GAACAG	ATG	ATG	1522					
Qy	810	TAG	TAG	AGAT	TC	CAAG	AAAT	TCTAT	AT	TG	AAAAA	CAAA	TAG	AG	869					
Db	1524	AGAA	GAC	CAAGAA	AAAAA	GAAGAA	CGACG	AAAGAA	CAAA	ATG	AG	AAAAA	CAAG	1583						
Qy	870	AAAT	TTT	GAAG	ACG	AAAGG	ATATCA	TAT	AAAGG	CATCT	TG	GA	AAAAA	GAAG	ATG	929				
Db	1584	AAAA	GAC	GAAG	AGG	GAACAG	AT	TAG	AG	AG	AT	TAG	AG	AT	GA	1644				
Qy	930	AATA	AAAC	CAAG	CAAA	AAAG	AGAT	TAA	TCTG	CTC	ACAT	CAAA	TCA	CTA	TAG	AG	989			
Db	1644	AAAA	AAAA	GTAG	AA	AAAAA	CAAG	AAAG	ACG	AAAG	AA	CTG	AA	GA	GA	GA	1703			
Qy	990	AGT	AAAT	TTT	CTG	ATG	TTT	CTAAT	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	1045			
Db	1704	AGAA	AC	AG	CTG	AAAAA	CAAG	AAAAA	CGAAG	GGG	AA	CGAT	TAT	GA	GA	1763				
Qy	1050	TGA	AT	TAT	GAC	AT	TAT	TAT	ATG	ATG	AAAG	ATG	ATG	AAAG	CTT	TAG	ACG	AT	1105	
Db	1764	AGAT	AT	CA	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	1822	
Qy	1110	GCT	AT	TG	TC	AT	TG	CA	AT	TTT	CAAG	ATG	AA	AAAA	CA	TAG	AA	CT	1165	
Db	1824	CAAA	GA	AAAA	CAAG	AAAG	CGG	AA	CGA	TG	ATG	AA	GA	CA	AA	CT	1883			
Qy	1170	AGA	AG	AT	TG	AT	TG	AAAAA	TTT	TG	AT	TG	AT	TG	AT	TG	AT	TG	1222	
Db	1884	AGAC	GA	CC	CAAG	AG	AG	CGG	AA	GA	AG	AT	TA	AG	AG	CG	CA	AA	GA	1943
Qy	1230	ATC	AG	AA	GA	AT	TAT	TG	AAAAA	AT	TAA	AAAAA	GA	AAAA	CA	AA	GA	1288		

Qy	1350	TGATPACAGGTTAATAAGGAAAAATTCATTAATCTTTTCATATATTGA	1409
Db	2061	AGAAAAAGACAAAGAAAGATTAACTAAGAAAGACAAAGATGACAAAGAAAAAGACAA	2120
Qy	1410	CGAGAGACAATCAAAATTTTACAGATCTGTGATGAGTTATCTGAAGATATA	1458
Db	2121	AGAGATGACAAAGAAAAAGACAAAGAAAGATTAACAAAGAAAAAGACAAA	2169
RESULT 15			
ID	AAA70212		
XX	AAA70212 standard; DNA; 1998 BP.		
AC	AAA70212;		
XX			
DT	07-NOV-2000 (first entry)		
XX			
DE	Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:345.		
XX			
KM	Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;		
XX	antimalarial; malaria; protozoacide; infection; insecticide; db.		
KW			
XX	Plasmodium falciparum.		
OS			
XX	WO200025728-A2.		
PN			
PD	11-MAY-2000.		
XX			
PF	05-NOV-1999; 99WO-US026796.		
XX			
PR	05-NOV-1998; 98US-0107131P.		
XX			
PA	(HOFF/) HOFFMAN S.		
PA	(CARU/) CARUCCI D.		
PA	(GARD/) GARDNER M.		
XX	(VENT/) VENTER J C.		
XX			
PI	Hoffman S, Carucci D, Gardner M, Venter JC;		
XX			
DR	WPI; 2000-365347/31.		
XX			
PT	Proteins encoded by chromosome 2 of the human malarial parasite.		
PT	Plasmodium falciparum, useful as antimalarial vaccines and in the		
PT	diagnosis of P.falciparum infection.		
XX			
PS	Disclosure; Page 537; 577pp; English.		
XX			
CC	The present invention describes proteins and their fragments (I) encoded		
CC	by chromosome 2 of the human malarial parasite, Plasmodium falciparum.		
CC	Also described are: (I) nucleotide sequences (II) encoding (I); and (2)		
CC	vaccines against P. falciparum infection comprising (I) or (II). (I) and		
CC	(II) are useful for the development of vaccines against P. falciparum		
CC	infection. (I) and polyclonal antisera or a monoclonal antibody raised to		
CC	immunogens comprising the sequences of (I), are useful in the detection		
CC	of infection with P. falciparum. Furthermore, (I) (especially when they		
CC	are rifins or secreted or membrane proteins) can aid the identification		
CC	of drugs to treat or prevent P. falciparum infection, or they can be used		
CC	to identify drug resistance in P. falciparum. Sequencing of the		
CC	Plasmodium chromosome 2 and the subsequent identification of proteins		
CC	encoded by it will help to expand our understanding of parasite biology,		
CC	a process hampered by the complexity of the parasite life cycle, and		
CC	provide new targets for vaccine and drug development. Parasite resistance		
CC	to drugs and mosquito resistance to insecticides have led to a resurgence		
CC	of malaria in many parts of the world, and there is a pressing need for		
CC	vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352		
CC	represent nucleotide and protein sequences given in the present		
CC	invention, but which are not specifically mentioned within the		
CC	specification		
XX			
XX			
SO	Sequence 1998 BP; 1127 A; 141 C; 338 G; 392 T; 0 U; 0 Other:		
Query Match	11.5%; Score 170.6; DB 3; Length 1998;		
Best Local Similarity	45.7%; Pred. NO. 1.2e-15;		

Matches	594;	Conservative	0;	Mismatches	706;	Indels	0;	Gaps	0;
Qy	80	ATAGACTTCTAAGAAAGTTTCAAGAGACGAAAGCGATTATTAACAAGAGACTTG	139						
Db	413	ATAGACAAAATGAAATTTTCAAGATGGAAGAAATGACAAAGCATTAATTTAAATG	472						
Qy	140	CTTAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGACGCGCTAAAGAA	199						
Db	473	AAAAAGAAAACGAGATATCTTAATTCATGCAAGAACAAAAGATGTAAGAAAG	532						
Qy	200	AGTTGCAAGAACAAAGCGATTTTGAACAAGAGACGCTGTAAAGAAAGTTGCAAG	259						
Db	533	GAGTAAAGATTTAGAAAGAAAAGAAAAGAAAATTTGATGATCATTAAGTGG	592						
Qy	260	AACAAACAAAGCGATTTAGAACAGATAGACTTCTAAAGAAAAGTTACAAAGCAGCAAA	319						
Db	593	AGGAAAACAAAATAATGATGACCATTAAGTTGAGGAAAACAAAATAATGATCATTA	652						
Qy	320	GGGATTTAGAACAAAGAGACGCTGTAAAGAAAAGTTGCAAGAACAAAGCGATTAG	379						
Db	653	AAGTTGAGGAAAACAAAATAATGATGATCATTAATAGAAAGATTAAGAAAAGTAGAG	712						
Qy	380	AACAAGAGACGCTGTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGA	439						
Db	713	AACATGAAGAGACGAAAGAGGATTAAGAAAGAAAATTCAGAAAACAAAATAAG	772						
Qy	440	GACTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGCTGCTA	499						
Db	773	ATGAATAATTAAGTTGAAAATGATGAATATGATGAATAATGATGTAAGTGAAGTTG	832						
Qy	500	AAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTGTAAAGAAAAT	559						
Db	833	ATGATGATTTGAGAAAGATTAATAATGAATAATGATGATTAATGATGATTAAGAAA	892						
Qy	560	TGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTGTAAAGAAAAGTTGCAAGAC	619						
Db	893	CAGATTAACATCTTTGGAAGAGAGGAAATGAATTAATGAAAAGCAATTTAGATGA	952						
Qy	620	ACCAAGAGATTTAGAACAAAGAGCGCTGATTCGAAAAAAATTTAGAAAGAAAAG	679						
Db	953	ACAAAAGATGTTAAATAATTAAGATCTAAGAAAGAAAAGTAAGATCTGAGAAAG	1012						
Qy	680	AACATGAGATTAATTAAGAGAGATTTATGCTGTTAGAAAACACAGCTATTAAGAC	739						
Db	1013	AAAAAGTAAGATTAAGAGAGGAAAGAAAAGTAAGATTAAGAGAGGAAAAAGTAAG	1072						
Qy	740	TTCATCAAGAAAATGAAAGCTGATTAATTAATTAATCAACATCAATCTTTACTCAGACA	799						
Db	1073	ATTAAGAGAGGAAAAGGTAAGATTAAGAGAGGAAAAGTAAGATTAAGAGAGG	1132						
Qy	800	ACAGAGGGAATAGTAGAGATTCCAAGGAAATCTTAATTAAGAAAACAAATAGAGAT	859						
Db	1133	AAAAAGAAAAGATTAAGATTAAGAGAGGAAAAGGTAAGATTAAGAGAGGAAAAG	1192						
Qy	860	CTATTACAAACAATGTTGAAGACGAGAGGATTAATCAATAAGACATCTTGAAGAAAAG	919						
Db	1193	AAAAAGATTAAGATTAAGAGAGGAAAAGGAAAGATTAAGAAAATTAAGAAATTAAC	1252						
Qy	920	AAGATGTTCAATTAAGAACAGAACAAAGAAAGATTAATCTCTGACATACAAAATCAT	979						
Db	1253	AAAAATGATTAAGAAAAGAGAGCGATTAAGAAAAGAAAAGTAATTAAGAGATTAAC	1312						
Qy	980	CATTAGAGACAGTAATTAATTTCTGATGTTAATGATTTCAAAATAGTAAGTGAAGATG	1039						
Db	1313	ATGATGATTAATGATGTAAGAAAATGATGAAGAAATTAAGAAAATGATGAGAAAG	1372						
Qy	1040	AAATAAGTCTGAATATGACGATTCATTAATGATGAAGAAAGAGATGTAAGACTTAG	1099						
Db	1373	ATGAGAGATTAAGCATGAGAAAACAAAAGAAAAGAAAAGAAAAGTAAGAAATGAAA	1432						
Qy	1100	ACGAATTTAAGCCTATTTGCAATATGACAAATTTCCAAAGATGAAGAAAACATAGCAATTT	1159						
Db	1433	ATGGAATTAAGAAATGGAAGTGAAGAAATGGAATTAAGAAATGGAATTAAGAA	1492						

Qy	1160	ATTAAGAACTAGAAAGTTTGATAGAGAAAATGAAAATTTAGATGATTTAGAGAGAA	1219
Db	1493	ATGAAAATTAAGAAATGAAAGTGAAGAAATGAAAAGGAAAATGAAAAGAAATGAA	1552
Qy	1220	TAGAAAATCATCAGAAAGATTAATTCGAGAGAAAATTAAGAAAAGGAAAGAAATATGAA	1279
Db	1553	ACGAAAATTAAGAAACGAAAAGGAAAAGGAAAAGTAAATTTAAAGAGATTGAAAAG	1612
Qy	1280	AAACAAAGATTAATTAATTTTAAACCAATGATTAAGGTTTGTATGATGAGCATATTTAA	1339
Db	1613	TAAACAATGCAACAAGGAAAATCTATGAAAAATTAATTAATTAATCTGAAAATTAACAATTA	1672
Qy	1340	AATTAATTAATGATTAAGACGTTTAATTAAGGAAAAGGAAA	1379
Db	1673	CAAAATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1712
RESULT 16			
AS46735/c			
ID	AAS46735 standard; DNA; 6292 BP.		
XX			
AC	AAS46735;		
XX			
DT	18-DEC-2001 (first entry)		
XX			
DE	Tumour suppressor gene derived chemically modified sequence #459.		
XX			
KW	Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;		
KW	tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;		
KW	cytosine methylation; de.		
OS	Homo sapiens.		
XX			
PN	WO200168912-A2.		
XX			
PD	20-SEP-2001.		
XX			
PF	15-MAR-2001; 2001WO-BP002955.		
XX			
PR	15-MAR-2000; 2000DE-01013847.		
XX			
PR	06-APR-2000; 2000DE-01019058.		
XX			
PR	07-APR-2000; 2000DE-01019173.		
XX			
PR	30-JUN-2000; 2000DE-01012529.		
XX			
PR	01-SEP-2000; 2000DE-01043826.		
XX			
PA	(EPIC-) EPIGENOMICS AG.		
XX			
PI	Olek A, Piepenbrock C, Berlin K;		
XX			
DR	WPI; 2001-602752/68.		
XX			
PT	Fragments of chemically modified genes associated with tumor suppressor		
PT	genes and oncogenes, useful in designing primers and probes for analyzing		
PT	diseases associated with cytosine methylation state e.g. cancer.		
PS	Claim 1; SEQ ID NO 459; 27pp; English.		
XX			
CC	The invention relates to a nucleic acid comprising a sequence of 18		
CC	bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with		
CC	bisulphite, of genes associated with tumour suppression and oncogenes		
CC	having a sequence taken from 536 (actually 533 since numbers 408, 458 and		
CC	500 are missing from the sequence listing) sequences (Ss) and sequences		
CC	complementary to (Ss). The nucleic acid may be a peptide nucleic acid-		
CC	oligomer (PNA) of at least 9 nucleotides and may form part of a set of		
CC	probes for detecting the cytosine methylation state and/or single		
CC	nucleotide polymorphisms and also to be used in an array for analysing		
CC	diseases associated with CpG dinucleotides e.g. cancers and tumours. The		
CC	probes can also be used in a method for ascertaining genetic and/or		
CC	epigenetic parameters for the diagnosis and/or therapy of existing		
CC	diseases or the predisposition to specific diseases, by analysing		
CC	cytosine methylations. The parameters may be compared to another set of		
CC	genetic and/or epigenetic parameters, the differences serving as basis		

CC for diagnosis and/or prognosis events which are disadvantageous to
 CC patients. The present sequence is one of the 533 genomic sequences
 CC derived from tumor suppressor genes and oncogenes. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 6292 BP; 1736 A; 99 C; 1304 G; 3153 T; 0 U; 0 Other;

Query Match 10.2%; Score 151.2; DB 4; Length 6292;
 Beel Local Similarity 46.2%; Pred. No. 6.1e-13;
 Matches 619; Conservative 0; Mismatches 708; Indels 13; Gaps 3;

```

QY 40 AAAGAAAGTTGCAAGAACCAAGACGATTAGAACCAAGATGAGCTTCTTAAGAAAG 99
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4548 AATATATATATATATATATATATATATATATATATATATATATATATATATAT 4489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 100 TTACAGAGCAGCAAGCGATTAGAACAGAGAGACTTGTAAAGAAAGTTGCAAGAA 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4488 CATATAAAACGATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 4429
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 160 CAACAAAGCGATCTAGAACAGAGAGCGTCTAAGAGAAAGTTGCAAGAACCAAGAGC 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4428 AAACAAATTAATTAACCAAAAAAAAAAAAAAAAAATTTGACAAAAAATTAATTAATA 4369
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 220 GATTAGAAACAAGAGAGAGCTGCTAAGAAAGTTGCAAGAACCAAGAGCTTTAGAA 279
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4368 TTTAATATATATATATATATATATATATATATATATATATATATATATATATAT 4309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 280 CAAGATGAGCTTCTTAAGAAAGTTTACAGAGCAGCAAGCGATTAGAAC---AAGA 335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4308 AAAAAAAAAATTTACAAAAATTAATTAATAATAATAATAATAATAATAATAATAATA 4249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 336 GAGAGCTGCTAAGAAAGTTGCAAGAACCAAGAGCTTTAGAACAGAGAGCGTGC 395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4248 AAAAAAAAAATTTAATATATATATATATATATATATATATATATATATATATATAT 4189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 336 TAAAGAAAGTTGCAAGAACCAAGAGCTTTAGAACAGAGAGCTTGTAAAGAAAA 455
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4188 CAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 456 GTTGCAGAGCAACAAGCGATTAGAACAGAGAGCGTCTAAGAAAGTTGCAAGAA 515
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4128 CAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4069
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 516 ACAACAAAGCGATTAGAAC---AAGAGAGCGTCTAAGAAAGTTGCAAGAACAC 571
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4068 AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4009
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 572 AAAGCGATTAGAACAGAGAGCGTCTAAGAAAGTTGCAAGAGCAGCAAGAGATT 631
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4008 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3949
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 632 TAGAAACAAGAGCGTGTATCGAAAAAATAATTAGAAAGAAAGAAACATGAGATA 691
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3948 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3889
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 692 TATTAGCAGAGATTATATGTCGTTAGAAATCCAGCTATATAACTTCATCAGAAA 751
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3888 TAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3829
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 752 ATGAACGTGATATATATATACACATCAATCTTCTTACCTCAGAGCAACAAGAGGATA 811
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3828 AACAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3769
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 812 GTAGAGATCCAGAGAAATATCTATATATAGAAACCAATATAGAGATCTATTAACA 871
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3768 CAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3709
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 872 ATGTGAAGAGCAGAGGATATATATATAGAGCATCTTGAAGAAAGAAAGTGTGTTCA 931
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3708 AACTTACAAAAAAGCAGCAAAAAAATATTTTAAATTAATTAATAATAATAATAATA 3649
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 932 TAAACCGAACAACAAAGAGATTAATCTGCTGACATACCAAAATCATACATTAGAGACG 991
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3648 CAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3589
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 992 TAAATATTTCTGATGTTAATGATTTTCAAAATAGTATAGAGATGAAATTAAGTGT- 1050
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3588 AAAAAATTAACGACCAACAAATTAATAATAATAATAATAATAATAATAATAATA 3529
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1051 ----GAATATGACATTCATTATATAGATGAAGAAAGAAAGATGATGAAGCTTAGACGATT 1106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3528 TTAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1107 TAACTTATTTGCAATATGACAAATTTCCAAATGAAAGAAACATAGGAATTTTAAGA 1166
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3468 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3409
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1167 ACTGAAGATTTGATTAAGAAAAATGAAAAATTTAGATGATTTAGATGAAGAAATGAAAA 1226
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3408 AATTAACGATTAATTAATTAATAATAATAATAATAATAATAATAATAATAATA 3349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1227 ATCATCAGAAAGATTAATCTGAAAGAAAAATTAATAATAATAATAATAATAATA 1286
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3348 AAAAAATTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1287 GATTAATTAATTTTAAACCAATGATTAAGTTGTATGATGAGCATATTAATAATA 1346
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3288 GAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1347 AATGATTAAGCAGGTTAATA 1366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3228 TAATTAATATTAATTTCAAAA 3209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 17
 AAS46608/C
 ID AAS46608 standard; DNA; 6767 BP.
 XX AAS46608;
 AC
 XX 18-DEC-2001 (first entry)
 DT
 XX Tumour suppressor gene derived chemically modified sequence #330.
 DE
 XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
 KW tumour; Opg dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 XX
 XX Homo sapiens.
 OS
 XX WO200168912-A2.
 PN
 XX 20-SEP-2001.
 PD
 XX 15-MAR-2001; 2001WO-EP002955.
 PF
 XX 15-MAR-2000; 2000DE-01013847.
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 PR
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K;
 PI
 XX WPI; 2001-602752/68.
 DR
 XX Fragments of chemically modified genes associated with tumor suppressor
 PT genes and oncogenes, useful in designing primers and probes for analyzing
 PT diseases associated with cytosine methylation state e.g. cancer.
 XX
 XX Claim 1; SEQ ID NO 330; 27bp; English.
 XX

CC The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pre-treated DNA (CP DNA) e.g. with
CC biolipidite, of genes associated with tumour suppression and oncogenes
CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and
CC 500 are missing from the sequence listing) sequences (Ss) and sequences
CC complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of
CC probes for detecting the cytosine methylation state and/or single
CC nucleotide polymorphisms and also to be used in an array for analysing
CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The
CC probes can also be used in a method for ascertaining genetic and/or
CC epigenetic parameters for the diagnosis and/or therapy of existing
CC diseases or the predisposition to specific diseases, by analysing
CC cytosine methylations. The parameters may be compared to another set of
CC genetic and/or epigenetic parameters, the differences serving as basis
CC for diagnosis and/or prognosis events which are disadvantageous to
CC patients. The present sequence is one of the 533 genomic sequences
CC derived from tumour suppressor genes and oncogenes. Sequences with even
CC numbered Seq ID numbers are the complementary sequence of the
CC corresponding odd numbered sequence (e.g. ID 2 and ID1, ID 536 and ID
CC 535, except for those whose partner sequence is missing). Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
CC
CC
CC

Sequence 6767 BP; 1446 A; 207 C; 1601 G; 3513 T; 0 U; 0 Other;

Query Match 10.1%; Score 149.2; DB 4; Length 6767;
Best Local Similarity 46.1%; Pred. No. 1.2e-12;
Matches 643; Conservative 0; Mismatches 743; Indels 8; Gaps 4;

QY 2 AAGAACCAACGACGATCTAGAACCAAGAGACGCTTAAGAAAGTTGCAAGAACAC 61
DB AATTAATTAACCAACGACGATCTAGAACCAAGAGACGCTTAAGAAAGTTGCAAGAACAC 3569
QY 62 AAAGGATTTAGAACAGATAGACTTGTCTAAGAAAGTTTACAGAGCAGCAAGCGATT 121
DB 3568 AAA--CATTAACACTCAAAATTTCCACTTAAATAATTAATAATCAATAATCTCCAA 3511
QY 122 TAGAACCAAGAGACTTGTCTAAGAAAGTTGCAAGAACCAAGCGATTGAGAACAG 181
DB 3510 AATAAAATTAATTAATACTTAACCTTAACCTTTCTTAATAAATCTAATAATTA 3451
QY 182 AGAGACGCTCTAAGAAAGTTGCAAGAACCAAGCGATTGAGAACAGAGACGCG 241
DB 3450 AAAACCCGAAAAATCTTAATATATAAAAACTTAATAAATCTTAATAAATTA 3391
QY 242 CTAAAGAAAGTTGCAAGAACCAAGCGATTGAGAACAGATGACTTGTCTAAGAA 301
DB 3390 CAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3331
QY 302 AGTTACAGAGACGACGATTTAGAACCAAGAGACGCTCTAAGAAAGTTGCAAG 361
DB 3330 ACAAATAAATCAATAAATCAATAAATCAATAAATCAATAAATCAATAAATCA 3271
QY 362 AACACCAAGCGATTGAGAACAGAGACGCTCTAAGAAAGTTGCAAGAACCA 421
DB 3270 ACATCAAAAAACAAAAAATACATAAATAAATAAATAAATAAATAAATAA 3212
QY 422 GCGATTTAGAACCAAGAGACTTGTCTAAGAAAGTTGCAAGAACCAAGCGATTGAG 481
DB 3211 AAAAATAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAA 3152
QY 482 AACAGAGAGACGCTCTAAGAAAGTTGCAAGAACCAAGCGATTGAGAACAGAG 541
DB 3151 AAATTAACCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 3092
QY 542 GACGCTCTAAGAAAGTTGCAAGAACCAAGCGATTGAGAACAGAGACGCTCTA 601
DB 3091 AACCAAAATTAACCAAAATTAATAATAATAATAATAATAATAATAATAATA 3032
QY 602 AAGAAAGTTGCAAGAGACGAGATTTAGAACCAAGAGCGTGTAGCAAAAAA 661

DB 3031 AAAAAAATATCAAAAAACAAAAACAAACAAATAATAACAAACCAAAAAATAAAAAA 2972
QY 662 ATTTAGAAAGAAAGAAACATGAGATATATTAGCAGAGATTTATATGTCGTTAG 721
DB 2971 TCAAAAAACAAAAAATCAAAAAACAAACATTAATAAATAAATAAATAAATAA 2912
QY 722 AAATACGACTTATAGACTTCCATCAGAAATGACGTCGATATATATATACCATCAAT 781
DB 2911 AACAAAAACAAAAAATCAAAAAACAAATAATAACAAACCAAAAAATAAAAAAT 2852
QY 782 CTTCTTTCCTCAGACACAGAGGAAATAGTAGATTCAGAGAAATATCTATAT-- 839
DB 2851 ATCAAAAAACAAAAAATCAAAAAACAAATAATAACAAATAAATAAATAAATAA 2792
QY 840 -AGAAAAACAAATAGAAATTTATTAACCAATGTTGAGAGAGAGATATACATA 898
DB 2791 AAAAAAATCAAAAAACAAACATTAATAAATAAATAAATAAATAAATAAATAA 2732
QY 899 AAGACATCTTGAAGAAAGAAAGATGTTCAATTAACAGAACAAAGAAAGATTAAT 958
DB 2731 AACAAAAACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2672
QY 959 CTGCTGACATACAAATCATACATTAGACAGTAATAATTTCTGATGTTATGATTTTC 1018
DB 2671 AAACCAAAACCAAAAAACAAAAAATCAACAACCAAAAAACAAACCAAAAAATA 2612
QY 1019 AATAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1078
DB 2611 AAAAAAATCAAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAA 2552
QY 1079 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1138
DB 2551 AAAAAAATCAAAAAACAAACATTAATAAATAAATAAATAAATAAATAAATAA 2492
QY 1139 ATGAAGAAACATAGAAATTTATTAAGAACTAGAAAGTTGATAGAAAGAAATTA 1198
DB 2491 AAAAAAATCAAAAAACAAAAAAATTAATAAATAAATAAATAAATAAATAAATAA 2432
QY 1199 TAGATGATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1258
DB 2431 AAAAAAATCAAAAAACAAAAAAATTAATAAATAAATAAATAAATAAATAAATAA 2372
QY 1259 AAAAAAAGAAATGAAAGAAACAAAGATTAATTTTAAACCAATGATTAAGT 1318
DB 2371 ACCAA--AAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2314
QY 1319 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1378
DB 2313 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2254
QY 1379 AATTCAATAAATCA 1392
DB 2253 AAAAAATTAATAAACA 2240

RESULT 18
ABL05274/C
ID ABL05274 standard; cDNA; 3662 BP.
XX
XX ABL05274;
DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 10304.
DB Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.


```

XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI: 2001-656860/75.
XX P-PSDB; ABB61171.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX Claim 1; SEQ ID NO 10304; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signaling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins (ABB57737-
XX ABB72022). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pcc_sequences
XX
XX Sequence 3662 BP; 661 A; 941 C; 421 G; 1639 T; 0 U; 0 Other;
SQ
Query Match 9.9%; Score 146; DB 4; Length 3662;
Best Local Similarity 46.7%; Pred. No. 3.5e-12;
Matches 603; Conservative 0; Mismatches 675; Indels 12; Gaps 4;
QY 24 ACAAGAGACGCTGCTAAAGAAAGTTGCAGAACACAAACGATTTAGAACAAAGATAG 83
DB 1859 AAAAAAGAGATGAATTAACAGAAATCAAGGAAACCAAGATTAATGAAAAGGAAAG 1800
QY 84 ACTTGCTAAAGAAAGTTTCAAGAGAGACAAAGCCGATTTTGAACAAGAGACTTGCTAA 143
DB 1799 A--ACTTAAAGAGAGAGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1742
QY 144 AGAAAGTTGCAAGAACAAACGATCTAGAACAAAGAGAGACGCTGCTAAAGAAAGTT 203
DB 1741 AGAGAAAGAGAAAGAAAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1682
QY 204 GCAGAGAACAAACGATTTTGAACACAGAGAGACGCTGCTAAAGAAAGTTGCAAGAA 263
DB 1681 GAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1626
QY 264 ACNAAAGCGATTTTGAACAAAGATAGACTTCTTAAAGAAAGTTTGAAGAGCCAGCA 323
DB 1625 GGAAAGAGAAAGAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1571
QY 324 TTTAGAACAAAGAGAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTTGAACA 383
DB 1570 GAAAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1511
QY 384 AGAGAGACGCTGCTAAAGAAAGTTGCAAGAACAAACGATTTTGAACAAGAGAGACT 443
DB 1510 AGAAAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAG 1451
QY 444 TGCTTAAAGAAAGTTGCAAGAACAAACGATTTTGAACAAGAGAGACGCTGCTTAAAG 503
DB 1450 GAAAGAGAAAGAGAGAGAAAGAAAGAAAGAAAGAAAGAGATGATGAGAAAGAG 1391
QY 504 AAAGTTGCAAGAACAAACGATTTTGAACAAGAGAGACGCTGCTTAAAGAAAGTTGCA 563
DB 1390 AGAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGAG 1331
QY 564 AGAACAAACAAAGCATTTTAGAACAAAGAGAGACGCTGCTTAAAGAAAGTTGCAAGAG 623

```

```

DB 1330 GAAAGAAAGAGAGAAAGGAAAGCAGAGAGAAAGAAAGAAAGAGAGAGAGAG 1271
QY 624 AAGAGATTTTAAACAAAGGAAAGCTGATATCAAGAAATTTTGAAGAAAGAAAGCA 683
DB 1270 GAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1212
QY 684 TGAGATATATTTAGACAGAGATTTATATATGCTGCTTGAAGAAATACAGCTATAGACTTC 743
DB 1211 AGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1152
QY 744 ATCAGAAATTAACGCTGATATTTATATCAATCAATCTTTTACTTCAGAGCAACAG 803
DB 1151 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1092
QY 804 AGGGAATGTAGATTCACAGAAATTTCTTATATGAAAGAAAGAAATAGAAATCTAT 863
DB 1091 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1032
QY 864 TACAACAAATTTGAGAGAGAGAGAGATATATATTAAGAGATCTTGAAGAAAGAAAG 923
DB 1031 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 972
QY 924 TGCTTCAATTAACCCAGAACAAAGAAAGATTAATCTGCTGACATACAAATCATACAT 983
DB 971 AGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 912
QY 984 AGAGACGATTAATTTTCTGATGTTTATGATTTTCAATTAAGTATGAGAGATGAAT 1043
DB 911 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 852
QY 1044 AAGTCTGAATATATGATCATTAATATGATGAAGAGAGAGATGAGAGATTTAGAG 1103
DB 851 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 792
QY 1104 ATTTAAGCTTATTTGTCATATATGACAAATTTCCAGATGAGAGAGAGAGAG 1163
DB 791 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732
QY 1164 AGAATAGAAATTTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1223
DB 731 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 672
QY 1224 AAAATCATCAGAGAGATTTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1283
DB 671 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 612
QY 1284 AAAGATATATATTTTAAACCAATGATTA 1313
DB 611 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 582

```

RESULT 19
AAS45347/C
ID AAS45347 standard; DNA; 9539 BP.

XX AAS45347;
XX
XX 18-DEC-2001 (first entry)
XX
XX Chemically pretreated complementary DNA associated with cell cycle #26.
XX
XX Cell cycle: human; Cpg dinucleotide; cytosine methylation; HIV; aging;
XX human immunodeficiency virus; neurodegenerative disorder; solid tumour;
XX graft-versus-host disease; glomerular disease; Levy body disease; cancer;
XX arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritis;
XX immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; db;
XX PCR primer.
XX
XX Homo sapiens.
XX
XX WO200168911-A2.
XX

PD 20-SEP-2001.
XX
XX 15-MAR-2001; 2001WO-EP002945.
XX
XX 15-MAR-2001; 2000DE-01013847.
PR 06-APR-2000; 2000DE-01019058.
PR 07-APR-2000; 2000DE-01019173.
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
PA (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2001-602751/68.
XX
XX
XX Designing primers and probes for analyzing diseases associated with
PT cytosine methylation state e.g. arthritis, cancer, aging,
PT arteriosclerosis comprising fragments of chemically modified genes
PT associated with cell cycle.
XX
XX
XX Claim 1; SEQ ID NO 52; 28pp; English.
XX
XX Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA
CC molecules associated with the cell cycle and specific PCR primers of the
CC invention. The sequences are useful for detecting the methylation state
CC of all CpG dinucleotides in a sequence and therefore for analyzing
CC associated diseases. By analyzing cytosine methylations in the pretreated
CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
CC of existing diseases or the predisposition to specific diseases can be
CC ascertained. The parameters may be compared to another set of genetic
CC and/or epigenetic parameters, the differences serving as basis for
CC diagnosis and/or prognosis events which are disadvantageous to patients.
CC The sequences of the invention are useful for the diagnosis and therapy
CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,
CC aging, glomerular disease, Lewy body disease, arthritis,
CC arteriosclerosis, solid tumours and cancers
XX
XX Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 U; 0 Other;
SQ
Query Match 9.4%; Score 139.4; DB 4; Length 9539;
Beet Local Similarity 45.6%; Pred. No. 2.8e-11;
Matches 607; Conservative 0; Mismatches 716; Indels 8; Gaps 3;
QY 72 AGAACAAGTAGACTTGTCTAAAGAAAAGTTACAGAGCAGCAAGCGATTGAAACAGA 131
DB 1343 AAAAAATAAAAAACTAAAAAAACCTTCTATCTATAACGCAACAAAAATAAAAAAA 1284
QY 132 GAGACTTGCTAAGAAAGTTGCCAGAACAAAGCGATCTAGAACAGAGACGTGC 191
DB 1283 AAAAAATCAAAACCGAAACCATTAATTTCTAAAAAATTAACAAATACAAACAA 1224
QY 192 TAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAGAGACGTCTAAAGAAA 251
DB 1223 TAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1164
QY 252 GTTGCAAGAACAAAGCGATTAGAACAGATAGACTTGTCTAAAGAAAAGTTACAGA 311
DB 1163 AAAAAAATAAACTTATC-CTACATACCATCCACCCCAAACTCTTAAAAA 1105
QY 312 GGAGCAAAAGCGATTAGAACAGAGAGAGTGTAAAGAAAAGTTGCAAGAACAGAG 371
DB 1104 AAAAAAATAAAAGCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1045
QY 372 CGATTAGAACAGAGACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAG 431
DB 1044 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 985
QY 432 ACAAGAGAGACTTCTTAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAG 491
DB 984 AAAAAAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 925
QY 492 ACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAGAGACGTCTAA 551

DB 924 AAAAAACGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 865
QY 552 AGAAAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGTGTCTAAAGAAAAGTT 611
DB 864 AATACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 805
QY 612 GCAAGAGCAGCAAGAGATTGTAACAAAGAAAGCTGTATAGAAAAAATTTAGAAAG 671
DB 804 AAAAAAATAAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 745
QY 672 AAAAAAGCAATGAGATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 731
DB 744 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 691
QY 732 TATAGAACTTCATGAGAAATGAACGTGATATTAATTAATTAATTAATTAATTAAT 791
DB 690 AATCGTACACTACCTCAACCTTAACAAACAAAAAACCCTACCTCAAAAAAATATA 631
QY 792 TCAAGAACAGAGGGAATGTAAGATTCCAAGAAATATCTTAATTAATTAATTAATTA 851
DB 630 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 571
QY 852 TAGAGATCTATTACAAATGTTGAAGAGAGAGATATTAATTAATTAATTAATTAAT 911
DB 570 AAAAAAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 511
QY 912 AGAAAAAGAAAGTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 971
DB 510 AAAAAAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 451
QY 972 AATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1031
DB 450 AAAAAAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 392
QY 1032 TGAGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1091
DB 391 AAAAAAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 332
QY 1092 AGACTTAGACGAATTTAAAGCCTATTGTGCAATATGACAAATTTCAAGATGAAGAAACAT 1151
DB 331 AAAAAAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 272
QY 1152 AGAAATTTTAAGAACTTAAGAAATTTGATAGAGAAAAATGAATTTAGATGATTAGA 1211
DB 271 AAAAAAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 212
QY 1212 TGAAGAAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1271
DB 211 AAAAAAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 152
QY 1272 ATATGAAAAAATAAGATTAATTAATTTAAACCAATGATTAATTTGATATGATGAGA 1331
DB 151 AAAAAAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 92
QY 1332 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1391
DB 91 AAAAAAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 32
QY 1392 ATTGTTCAATA 1402
DB 31 TATATCTAAAA 21
RESULT 20
ABK28180/C
ID ABK28180 standard; DNA; 9539 BP.
XX
XX ABK28180;
AC
XX
XX 23-APR-2002 (first entry)
DT
XX
XX DNA transcription associated complementary genomic DNA #27.

XX DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
KM PNA: cytosine methylation state; SNP: retroviral infection; gene; ds;
KM single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
KM viral infection; Sezary syndrome; haematological disorder; tuberculosis;
KM immunological disorder; Werner syndrome; developmental disorder;
KM psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
KM neurodegenerative disorder; Mardenburg syndrome; Niemann-Pick disease;
KM myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
KM angiogenesis; congenital heart disease; HDR syndrome; gene therapy;
KM polyglutamine disorder; solid tumour.
XX Unidentified.
OS WO200192565-A2.
XX
XX
XX 06-DEC-2001.
XX
XX 06-APR-2001; 2001WO-EP003973.
XX
XX 06-APR-2000; 2000DE-01019058.
XX 07-APR-2000; 2000DE-01019173.
XX 30-JUN-2000; 2000DE-01032529.
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
PI WPI; 2002-090046/12.
XX
XX New nucleic acids or oligomers, useful for diagnosing or treating
PT diseases associated with DNA transcription, e.g. immunological disorders,
PT Werner syndrome, psoriasis, myocardial infarction, solid tumors or
PT cancer.
XX
XX Claim 1; SEQ ID NO 54; 32bp; English.
XX
XX The invention relates to a nucleic acid, which comprises a segment of the
CC chemically pretreated DNA of genes associated with DNA transcription from
CC one of 346 sequences, and an oligomer, in particular an oligonucleotide
CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
CC to the chemically pretreated DNA of genes associated with DNA
CC transcription. The set of oligomer probes are useful for detecting the
CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
CC in a chemically pretreated genomic DNA. The nucleic acids are useful for
CC diagnosing or treating diseases associated with DNA transcription
CC (particularly with the methylation status), e.g. adenosine deaminase
CC deficiency, viral infection, retroviral infection, Sezary syndrome,
CC haematological disorders, immunological disorders, Werner syndrome,
CC tuberculosis, developmental disorder, psoriasis, Rieger's syndrome,
CC neurological disorders, neurodegenerative disorders, Mardenburg
CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
CC infarction, hypertension, angiogenesis, erythropoiesis, congenital heart
CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
CC or cancer. Sequences ABK28127-ABK28472 represent DNA transcription
CC associated genomic DNA molecules of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification but
CC was obtained in electronic format directly from the European Patent
CC Office
XX
SQ Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 U; 0 Other;

Query Match 9 4%; Score 139.4; DB 6; Length 9539;
Beet local similarity 45.6%; Pred. No. 2.8e-11;
Matches 607; Conservative 0; Mismatches 716; Indels 8; Gaps 3;

QY 72 AGAACAATGACCTGCTAAAGAAAGTTACAGAGAGCAAAAGCGATTTAGAACAGCA 131
Db 1343 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1284
QY 132 GAGACTTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGAGAGCGTGC 191

Db 1283 AAAATCAAAACCCGAAACCATTAATTTCTTAATAATTAATTAATAATTAATTAACAA 1224
QY 192 TAAAGAAAGTGGAGAGACAAAGCAAGATTAGAACAGAGAGAGCTGTTAAGAAA 251
Db 1223 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1164
QY 252 GTTCCAGAAACAAAGAGCGATTAGAACAGATAGACTTCTTAAGAAAGATTAGACA 311
Db 1163 AAAACAAAACCTTAATTC-CTAATACATCCACCTCCAGAACTCTTCTTAATAA 1105
QY 312 GCAGCAAGCGATTAGAACAGAGAGAGCGTCTTAAGAAAGTTGCAAGAACAAAG 371
Db 1104 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1045
QY 372 CGATTTGAACAGAGAGAGCTGCTTAAGAAAGTTGCAAGAACAAAGCGATTTGA 431
Db 1044 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 985
QY 432 ACAAGAGAGACTTGCTTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAG 491
Db 984 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 925
QY 492 AGTGTCTTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGTCTTA 551
Db 924 AAAAACCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 865
QY 552 AGAAGAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGAGCGTCTTAAGAAAGT 611
Db 864 AAATCATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 805
QY 612 GCAAGAGAGAGAGAGATTAGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 671
Db 804 ACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 745
QY 672 AAAAAG 731
Db 744 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 691
QY 732 TATGAAGCTTCCATCAAGAAATGAAGAGATTAATTAATTAATTAATTAATTAATTA 791
Db 690 AATGTATCACTTACATCTCAACCTTAACAAAGAGAGAGAGAGAGAGAGAGAGAG 631
QY 792 TCAG 851
Db 630 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 571
QY 852 TAGGAATCTATTAACAATATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 911
Db 570 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 511
QY 912 AGAAG 971
Db 510 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 451
QY 972 AAATCATATTAAG 1031
Db 450 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 392
QY 1032 TGAG 1091
Db 391 ACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 332
QY 1092 AGACTTGAAG 1151
Db 331 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 272
QY 1152 AGAATTTAATTAAG 1211
Db 271 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 212
QY 1212 TGAAG 1271
Db 211 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 152

Oy		1272	TATGTAAGAAAAACAAAGCATATAATTATTTTAAACCAAATGATATAACTTGATGTAGCA	1331
Dd		151	AAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAACTAAAATAATTAATAAT	92
Oy		1332	TATTAAAAAATATTAATAATGATTAAGCAGTTATATAGAAGAAAGAAAAATTCATTAATC	1391
Dd		91	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAATTAATAAAAAAAAAAAAAAAAAAATTTTC	32
Oy		1392	ATTGTTGCATA	1402
Dd		31	TATATCTAAAA	21
 RESULT 21 ABLJ33697/C ID ABLJ33697 standard; DNA; 6668 BP. XX XX XX ABLJ33697; XX XX DT 26-MAR-2002 (first entry) XX DE Human immune system associated gene SEQ ID NO: 1670. XX XX KW Human; immune system disease; cytosine methylation; antiaesthetic; KW antiarteriosclerotic; antiamebic; cytostatic; nootropic; KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological; KW antirheumatic; antiarthritic; antidiabetic; arteriosclerotic; anaemia; KW antinflammatory; cancer; eye disease; arteriosclerosis; AIDS; KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy; KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene; de. XX XX Homo sapiens. XX OS XX WO200200928-A2. XX EN XX PD 03-JAN-2002. XX PP PP 02-JUL-2001; 2001WO-EP007537. XX PR 30-JUN-2000; 2000DE-01032529. PR 01-SEP-2000; 2000DE-01043826. XX PR XX PA (BPIG-) EPIGENOMICS AG. XX PI Olek A, Piepenbrock C, Berlin K; DR WPI; 2002-130909/17. XX PT Nucleic acid comprising fragment of chemically modified gene, useful for PT diagnosis and treatment of diseases associated with abnormal cytosine XX methylation. XX XX Claim 1; SEQ ID NO 1670; 32pp + Sequence Listing; German. The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention				

[illegible]

```
Oy 1164 AGAACTAGAGATTGTGATGAGAAAAATGAAATTTTGAATGATTTAGATGAGAAATGA 1223
Db 2721 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2662
Oy 1224 AAATCATCGAGAGATTATCTGAAAGAAAAATTAATAAGAAAGAAATATGAAAAAC 1283
Db 2661 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2602
Oy 1284 AAAGATATATATTTTAAACCAATGATAAAGTTGATGATGACATATTAATAATA 1343
Db 2601 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2542
Oy 1344 TAAAAATGATAGCAGCTTAATTAAGAAAGAAAAATTCATTAATCATGTTTCATAT 1403
Db 2541 AAATTAATCTATTTCTTAATAACGTACTAATTCATTAATAAAAAAAAAACCAATTTTATATA 2482
Oy 1404 ATTTGAC 1410
Db 2481 AATTAAC 2475

RESULT 22
AAK33181
ID AAK33181 standard; DNA; 6644 BP.
XX
XX AAK33181;
XX
XX 25-JUN-1999 (first entry)
XX
XX Base sequence of the plasmid pRx-lres-bar.
XX
XX Cowpox virus; bar; viral vector; expression; apoptosis; resistance; crmA;
XX autoclinic disease; graft rejection reaction; inflammation;
XX inflammatory disease; ss.
XX
XX Synthetic.
XX Cowpox virus.
XX
XX WO9913073-A2.
XX
XX 18-MAR-1999.
XX
XX 07-SEP-1998; 98MO-JP004010.
XX
XX 08-SEP-1997; 97JP-00259235.
XX
XX (RPRG-) RPR GENCELL ASIA PACIFIC INC.
XX
XX Hamada H;
XX
XX WPI; 1999-243728/20.
XX
XX New apoptosis-resistant virus-sensitive cell.
XX
XX Example 1; Page 38-41; 51pp; English.
XX
XX The present invention describes an apoptosis-resistant virus-sensitive
XX cell line into which an apoptosis resistance gene has been introduced.
XX The recombinant viruses generated are capable of expressing apoptosis-
XX associated genes. These can then be used in a variety of diseases for
XX which the induction of apoptosis by gene transfer, or where the
XX inhibition of harmful apoptosis, is therapeutic. The recombinant viruses
XX are useful as vectors for gene therapy which can be applied to cancer
XX therapy for destroying cancer cells selectively, the treatment of
XX autoimmune diseases and graft rejection reaction, and apoptosis induction
XX therapy for inflammatory cells in inflammatory diseases. Prior arts have
XX encountered the problem where if an adenovirus vector capable of
XX expressing an apoptosis-associated gene is introduced into animal cells,
XX the cells producing the virus will be destroyed because the period of
XX time required to induce cell death by apoptosis is shorter than that
XX required to replicate and produce the virus, resulting in failure to
```

```
CC obtain a recombinant virus having the integrated apoptosis-associated
CC gene. In this invention an apoptosis-resistant 293 cell line (having an
CC apoptosis resistant gene introduced) is established and overcomes the
CC problem. The present sequence represents the base sequence of the plasmid
CC pRx-lres-bar, which contains the cowpox virus bar gene, and is used in an
CC example from the present invention
XX
XX Sequence 6644 BP; 2166 A; 1573 C; 1424 G; 1481 T; 0 U; 0 Other;
SQ
Query Match 9.2%; Score 136.4; DB 2; Length 6644;
Best Local Similarity 49.6%; Pred. No. 7.7e-11;
Matches 350; Conservative 0; Mismatches 356; Indels 0; Gaps 0;
Oy 1 CAAGACACCAACGATCTCTGAACAGAGAGAGCTGCTTAAGAAAAGTTGCAAGACAA 60
Db 3738 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3797
Oy 61 CAAAGCGATTTAGAACAGATAGACTTGCTTAAGAAAAGTTTACAGACAGCAAGCGAT 120
Db 3798 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3857
Oy 121 TTGAAACAGAGAGACTTGCTTAAGAAAAGTTGCAAGACACAAAGCGATCTAGACAA 180
Db 3858 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3917
Oy 181 GAGAGAGCTGCTTAAGAAAAGTTGCAAGACACAAAGCGATTTAGAACAGAGAGAGCT 240
Db 3918 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3977
Oy 241 GCTTAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGATAGACTTGCTTAAGAAA 300
Db 3978 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4037
Oy 301 AAGTTCAAGAGCGCAAGCGATTTAGAACAGAGAGAGCTGCTTAAGAAAAGTTGCA 360
Db 4038 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4097
Oy 361 GAACACAAAGCGATTTAGAACAGAGAGAGCTGCTTAAGAAAAGTTGCAAGACAA 420
Db 4098 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4157
Oy 421 AGCGATTTAGAACAGAGAGACTTGCTTAAGAAAAGTTGCAAGACAAAGCGATTTA 480
Db 4158 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4217
Oy 481 GAACAGAGAGAGCTGCTTAAGAAAAGTTGCAAGACAAAGCGATTTGACAGAGAG 540
Db 4218 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4277
Oy 541 AGACGTGCTTAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGAGAGAGCTGCT 600
Db 4278 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4337
Oy 601 AAAGAAAAGTTGCAAGAGCAAGAGATTTAGAACAAAGAAAGCTGATACGAAAAA 660
Db 4338 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4397
Oy 661 AATTTAGAAAAGAAAAGAAATGAGATGATATTTAGCAGAGATT 706
Db 4398 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGATTGTCGAAGCTT 4443

RESULT 23
AAK33182
ID AAK33182 standard; DNA; 7372 BP.
XX
XX AAK33182;
XX
XX 25-JUN-1999 (first entry)
XX
XX Base sequence of the plasmid pRx-Bcl-xl-bar.
XX
XX Cowpox virus; bar; viral vector; expression; apoptosis; resistance; crmA;
```


Qy	241	GCTAAGAAAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGATAGACTTGCTTAAGAA	300
Db	5330	AA	5389
Qy	301	AAAGTTACAAGACGACAAAGCGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAA	360
Db	5390	AA	5449
Qy	361	GAACACAAACGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAAACAAACAA	420
Db	5450	AA	5509
Qy	421	AGCGATTTAGAACAAAGAGACTTGCTTAAGAAAGTTGCAAAGAACAAACGCGATTTA	480
Db	5510	AA	5569
Qy	481	GAACAGAGACGTCCTTAAGAAAGTTGCAAGAACAAACGCGATTTAGAACAAAG	540
Db	5570	AA	5629
Qy	541	AGACGCTGAAGAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAGACGTCCT	600
Db	5630	AA	5689
Qy	601	AAAGAAAGTTGCAAGACGACAAAGAGATTTAGAACAAAGAAAGCTGATATCGAAAAA	660
Db	5690	AA	5749
Qy	661	AAATTAGAAAGAAAAAGAACATGAGATATTTAGACAGAGATT	706
Db	5750	AAAGAAATTGTCGAAGCTT	5795

XX	RESULT 26
XX	ADSB89675/c
ID	ADSB89675 standard; DNA; 5493 BP.
XX	
XX	ADSB89675;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	Oligonucleotide of the invention SEQ ID NO:691.
XX	
KM	es; cell proliferative disorder; breast; methylation; cytostatic;
XX	gene therapy; single nucleotide polymorphism; SNP.
XX	
OS	Unidentified.
XX	
PN	WO2004035803-A2.
XX	
PD	29-APR-2004.
XX	
PF	01-OCT-2003; 2003WO-EP010881.
XX	
PR	01-OCT-2002; 2002DE-01045779.
PR	07-JAN-2003; 2003DE-01000096.
PR	17-APR-2003; 2003DE-01017955.
XX	
PA	(EPIG-) EPIGENOMICS AG.
XX	
PI	Foekens J, Harbeck N, Koenig T, Maier S, Martens J, Model F;
PI	Nimmrich I, Rujan T, Schmitt A, Schmitt M, Look MP, Marx A;
XX	
DR	WPI; 2004-348468/32.
XX	
XX	Predicting responsiveness of a subject with breast cell proliferative
PT	disorder, useful for treating or differentiating breast cell
PT	proliferative disorders comprises analyzing methylation pattern of a
XX	genomic DNA from the subject.
PS	
XX	Claim 25; SEQ ID NO 691; 104bp; English.
XX	

CC The invention relates to a novel method for predicting the responsiveness
CC of a subject with a cell proliferative disorder of the breast tissues to
CC a therapy comprising analysing the methylation pattern of a target
CC nucleic acid by contacting at least one of the target nucleic acids in a
CC biological sample obtained from the subject prior to or during treatment.
CC The method of the invention has cytosinatic activity, and may have a use
CC in gene therapy. The set of oligonucleotides comprising at least two of
CC the oligomers are useful for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms (SNPs) within the sequences. The
CC methods, nucleic acid, oligonucleotide, and kit are useful for the
CC treatment, characterisation, classification and/or differentiation, of
CC breast cell proliferative disorders. The method is also useful for
CC predicting the responsiveness of a subject with a cell proliferative
CC disorder of the breast tissues to a therapy. The present sequence is used
CC in the exemplification of the invention.

XX Sequence 5493 BP, 851 A, 0 C, 1692 G, 2950 T, 0 U, 0 Other;

Query Match	9.1%	Score 134.2	DB 13	Length 5493
Best Local Similarity	47.2%	Pred. No. 1.6e-10		
Matches 504	Conservative 0	Mismatches 558	Indels 5	Gaps 3
Qy	359	AAAGAACACAAAGGATTTAGAACACAGAGACGCTGCTAAAGAAAGTTGCAGAACAC	418	
Dy	5473	AAACCAAAAAAAAAAAAAAAAAACCAACAAAAACAAAAACAAAAACAAAAATAAA	5414	
Qy	419	AAAGCGATTTAGAACAGAGAGCTTGCTAAAGAAAGTTGCAGAACACAAAGCGATT	478	
Dy	5413	AAACACACAAAAAAAATACAAACAAAAACAAACACAAATAAAAAACAAAAAAC	5354	
Qy	479	TAGAACAGAGAGAGCGTGTAAAGAAAGTTGCAGAACACAAAGCGATTTAGAACAG	538	
Dy	5353	AATTAATAATTAATAACAAAAAACAACAAAAACAAAAAATTAATAA	5294	
Qy	539	AGAGACGTGCTAAAGAAAGTTGCAGAACACAAAGCGATTTAGAACAGAGACGTG	598	
Dy	5293	AAAAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	5234	
Qy	599	CTAAAGAAAGTTGCAGAGACAAAGATTTAGAACAAAGAGGCTGATACGAAA	658	
Dy	5233	AAAAATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	5174	
Qy	659	AAAAATAGAAAGAAAGAAAGCATGAGATATATTAGCAGAGATTTATATGTCGTT	718	
Dy	5173	AAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	5114	
Qy	719	TAGAAATACGAGCTATAGAACTTCATCAGAAAAAGACGTGATTTATATACCATC	778	
Dy	5113	CAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	5054	
Qy	779	A-ATCTTCTTACTCAGGACACAGGGGAATGTGAGATTCACAGAAATTCATTA	837	
Dy	5053	ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4994	
Qy	838	ATAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	897	
Dy	4993	TACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4934	
Qy	898	AAAGGACATCTTGAGAGAAAGAAAGATGTTCAATAAACCAGAACAAAAAGAGATAA	957	
Dy	4933	AAATAAAAATCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4874	
Qy	958	TCGTCGACATCAAAATCATCTTGAAGACGTAAATTTCTGATGTTAAGATTTT	1017	
Dy	4873	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4817	
Qy	1018	CAATTAAGTATGATGAGATGAATTAAGTGTGAATATGACATTCATTAATGATGAA	1077	
Dy	4816	CAAAATACAAAAACAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4757	
Qy	1078	GAAGAGATGATGAGACTTAGACGAATTTAAGCTTATGTCATATGACAAATTTCAA	1137	
Dy	4756	ATACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4697	

DT	12-JUL-2002	(first entry)
DE	Oligonucleotide for detecting cytosine methylation SEQ ID NO 2180.	
XX		
XX	Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;	
XX	drug; side effect; cancer; central nervous system; cardiovascular;	
KW	gastrointestinal; respiratory system; single nucleotide polymorphism;	
KW	SNP; cell differentiation; ds.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200218632-A2.	
XX		
PD	07-MAR-2002.	
XX		
XX	01-SEP-2001; 2001WO-EP010074.	
PF		
XX	01-SEP-2000; 2000DE-01043826.	
PR		
XX	05-SEP-2000; 2000DE-01044543.	
PR		
XX	(EPIG-) EPIGENOMICS AG.	
PA		
XX		
P1	Olek A, Piepenbrock C, Berlin K, Gietig D;	
XX		
DR	WPI; 2002-371829/40.	
XX		
PT	Determining the degree of cytosine methylation in genomic DNA, useful for	
PT	diagnosis and prognosis, comprises selective hybridization of amplicons	
PT	from chemically treated DNA.	
XX		
PS	Claim 12; 56pp + Sequence Listing; 56pp; German.	
XX		
CC	This invention describes a novel method for determining the degree of	
CC	methylation of a particular cytosine in a motif 5'-CpG-3', present in a	
CC	genomic sample of DNA. The sample is treated chemically to convert	
CC	cytosine (C) but not methylated C, to uracil, then part of the genomic	
CC	DNA that contains the target C is amplified to form a labeled amplicon.	
CC	The amplicon is hybridized to two classes, each with at least one member,	
CC	of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the	
CC	degree of hybridization to both classes is determined from the label on	
CC	the amplicon. From the ratio of labels hybridized to the two classes of	
CC	oligomers, the degree of methylation is calculated. The method is used:	
CC	(1) for diagnosis and/or prognosis of side effects of therapeutic drugs	
CC	and of a wide range of diseases, e.g. cancer, disorders of the central	
CC	nervous, cardiovascular, gastrointestinal and respiratory systems etc.,	
CC	particularly by detecting mutations or single nucleotide polymorphisms	
CC	(SNP's) and (ii) for differentiation of cell or tissue types and for	
CC	investigating cell differentiation. The method allows the methylation	
CC	status of many C residues to be determined simultaneously. ABQ13410-	
CC	ABQ54121 represent genomic DNA sequences used to illustrate the method	
CC	for determining the degree of cytosine methylation described in the	
CC	disclosure of the invention	
XX		
XX		
SO	Sequence 778 BP; 614 A; 29 C; 24 G; 111 T; 0 U; 0 Other;	
XX		
Query Match	9.0%;	Score 135.8; DB 6; Length 778;
Best Local Similarity	48.5%;	Pred. No. 2.9e-10;
Matches 365; Conservative	0;	Mismatches 387; Indels 0; Gaps 0
0Y	638	AAAGGAGCGTCGATCGAAGAAAAATTTTGAAGAAAGAAAGGACATGAGATATATTG 697
Db	6	AAAAAAAAAAATTCAGAAAAAAAATTTACTAATTAATAAAAAAATTTAATAAAAAA 65
0Y	698	CAGAGAGATTTTATGTCGTCGTTAGAAATCCAGACTATGACCTTCATCAGAAAAATGAC 757
Db	66	TTACGAAAAAAAATTTTCGAAAAAAAATTTATTAATAAAAAAATTTATTA 125
0Y	758	GTGATATTATATATACACATCAATCTTCTTTACCTCAGACAAAGAGGAATAGTAGAG 817
Db	126	AAAAAAAAAAATTCGAAAAAAAATTTATTAATAAAAAAATTTAATAAAAAA 185
0Y	818	ATTCCAGAGAAATATCTATATATGAAAAAAACAAATATGAGATCTATTACAAATGTTG 877

Db	186	ATTACGAAAAAAAAAAATTTATAAAAAAAAAAAAAAAAATTTATAAAAAAAAAAAAAAAAATTTACG	245
Qy	878	AAGGACGAGGGAATATACATTAAGGACATCTTGAGAGAGAGAGAGATGTTCAATATAAC	937
Db	246	AAAAAAAAAAAAAAAAATTTATAAAAAAAAAAAAAAAAATTTATAAAAAAAAAAAAAAAAATTTACGAAAAA	305
Qy	938	CAGAACAAAAAGATTAATCTGCTGACATCAAAAATCTACATTGAGACGTAAATA	997
Db	306	AAAAAATTTACGAAAAAAAAAAAAAAAAATTTACGAAAAAAAAAAAAAAAAATTTACGAAAAAAAAA	365
Qy	998	TTTCGATGTTAATGATTTTCAATTAAGTAAGTATGAGATGAAATGAGTGAATATG	1057
Db	366	TTACAAAAACGAAAAATTTATAAAAAAAAATTTATCGAAAAAAAAAAAAAAAAATTTACGAAAAA	425
Qy	1058	ACGATTCATTAATGATGAGAGAGAGAGATGATGAGACTTACGCAATTTTAAGCTTATG	1117
Db	426	AAAAAAAAAATTTACGAAAAAAAAAAAAAAAAATTTATAAAAAAAAAAAAAAAAATTTACGAAAAAAAAAAT	485
Qy	1118	TGCAATATGACATTTCCAGATGAGAGAGAAACATGGAATTTATTAAGAACTGAGAGATT	1177
Db	486	TACGAAAAAAAAAAAAAAAAATTTATAAAAAAAAAAAAAAAAATTTACGAAAAAAAAAAATTTACGAAAAA	545
Qy	1178	TGATAGAGAAAAATGAAAAATTTAGATGATTTAGTATGAGATGAGAAATTCATCAGAG	1237
Db	546	AAAAAAAAATTTATAAAAAAAAAAAAAAAAATTTACGAAAAAAAAAAAAAAAAATTTACGAAAAAAAAA	605
Qy	1238	AATATCTGAGAGAAAAATTAAGAGAGAGAAATATGAGAGAGAAACAGAGATTAATTT	1297
Db	606	AAATTTACGAAAAAAAAAAAAAAAAATTTACGAAAAAAAAAAAAAAAAATTTACGAAAAAAAAAATTT	665
Qy	1298	TTAAACCAATGATTAAGTTTGTATGATGACATATTAATAAATAATTAATAATGATTAAGC	1357
Db	666	ATTAATAAAAAAAAAAAATTTACGAAAAAAAAAAAAAAAAATTTATTAATAAAAAAAAAAAATTTATTA	725
Qy	1358	AGTTAATTAAGAGAAAGAGAAAAATTCATATAA	1389
Db	726	AAAAAAAAAATTTATTAATAAAAAAAAAAAATTTACGA	757

RESULT 29

ADS89401/c

ID ADS89401 standard; DNA; 5493 BP.

XX AC ADS89401;

XX DT 18-NOV-2004 (first entry)

XX DE Oligonucleotide of the invention SEQ ID NO:417.

XX KM ss; cell proliferative disorder; breast; methylation; cytostatic;

XX KM gene therapy; single nucleotide polymorphism; SNP.

XX OS Unidentified.

XX PN WO2004035803-A2.

XX PD 29-APR-2004.

XX PF 01-OCT-2003; 2003WO-BP010881.

XX PR 01-OCT-2002; 2002DB-01045779.

XX PR 07-JAN-2003; 2003DE-0100096.

XX PR 17-APR-2003; 2003DE-01017955.

XX PA (EPiG-) EPIGENOMICS AG.

XX PI Foekens J, Harbeck N, Koenig T, Maier S, Martens J, Model F;

XX PI Nimmrich I, Rujan T, Schmitt A, Schmitt M, Look MP, Marx A;

XX DR WPI; 2004-348468/32.

XX PT Predicting responsiveness of a subject with breast cell proliferative

XX disorder, useful for treating or differentiating breast cell

Db 2419 TAAAAAAATAAAAATCTAAAAAAATTAATAATCTAAAAAAATTAACATATTA 2360

Search completed: December 20, 2005, 05:54:22
Job time : 1151.77 secs

THIS PAGE BLANK (USPTO)